ihealth | Millennium Institute for Intelligent Healthcare Engineering

CENIA
CENTRO NACIONAL DE INTELIGENCIA ARTIFICIAL

A gentle tutorial on NLG in Medical Applications

https://github.com/PabloMessina/Code-for-SIPAIM-2022-NLG-Tutorial

Pablo Messina and Denis Parra
CS Departament, PUC Chile (also at CENIA & IMFD)
SIPAIM 2022
November 8th, 2022

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Acknowledgments

These slides thanks to:

- Jorge Perez Rojas, Founder and CTO of cero.ai

- Online draft of “Speech and Language Processing” (3rd ed.) by Jurafsky and Martin https://web.stanford.edu/~jurafsky/slp3/

ihealth | Millennium Institute for Intelligent Healthcare Engineering

NLP and NLG

Meanwhile **NLP** (Natural Language Processing) deals with many tasks related to representing natural language in order to process it with machines (sentiment analysis, entity linking, POS, etc.), **NLG** focuses on the specific task of **Generating Natural Language (text)**.

ihealth | Millennium Institute for Intelligent Healthcare Engineering

NLG in medical activities

NLG can then support physicians in medical activities such as:

- Writing radiological reports from images,
- Answering questions related to medical imaging,
- Rephrasing clinical text into a different styles,
- Summarizing findings into more concise text,
- etc.

ihealth | Millennium Institute for Intelligent Healthcare Engineering

NLG in iHEALTH

Currently, in iHEALTH we are researching NLG in the context of report generation:

- Chest X-ray report (Parra, Besa, Dunstan, et al.)
- Mammography report (Chabert, Salas, et al.)
- MRI prostate report (Besa, Parra, Caprile, et al.)

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Fundamentals: How do we generate text?

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Language Models

A Language model (LM) is a model which assign probabilities to sequences of words.

The simplest language model is the n-gram, a sequence of n words.

ihealth | Millennium Institute for Intelligent Healthcare Engineering

N-gram example

For instance, in a n-gram language model we could calculate the probability of a word (e.g. inflated) given some context $P(w|c)$:

$$P(\textit{inflated}|\textit{lungs are normally}) = \frac{C(lungs\ are\ normally\ inflated)}{C(lungs\ are\ normally)}$$

ihealth | Millennium Institute for Intelligent Healthcare Engineering

The cost of calculating probabilities

With a large corpus, we can calculate probabilities for many sequences $P(w1, w2, \ldots, wn)$ with different length “n”.

However, the longest the sequence (the “n” in n-gram) the more expensive is to calculate the probabilities and the data needed !

Relaxations to calculate LM probabilities

Two tricks are used to calculate the probability of sequences in LMs:

- The chain rule of probability

$$
\begin{aligned}
P(w_{1:n}) &= P(w_1)P(w_2|w_1)P(w_3|w_{1:2})\ldots P(w_n|w_{1:n-1}) \\
&= \prod_{k=1}^{n} P(w_k|w_{1:k-1})
\end{aligned}
$$

- Approximating by shortening the “history” (just a few words)

ihealth | Millennium Institute for Intelligent Healthcare Engineering

What is the state of the art in NLG (and why)

ihealth | Millennium Institute for Intelligent Healthcare Engineering

The state of the art in NLG

Neural LMs are nowadays the state of the art in NLG

The reason is that, unlike n-gram LMs, they can deal with unobserved sequences

Also, in the newest generation of generative LM (Transformers), many computations can be conducted in parallel

Example for Neural LM

Training data:
We've seen: granuloma within the **rigth** lung
Never seen: granuloma within the **left** lung
Test data: granuloma within the left ___

Example inspired by https://web.stanford.edu/~jurafsky/slp3/slides/7_NN_Apr_28_2021.pdf

Example for Neural LM

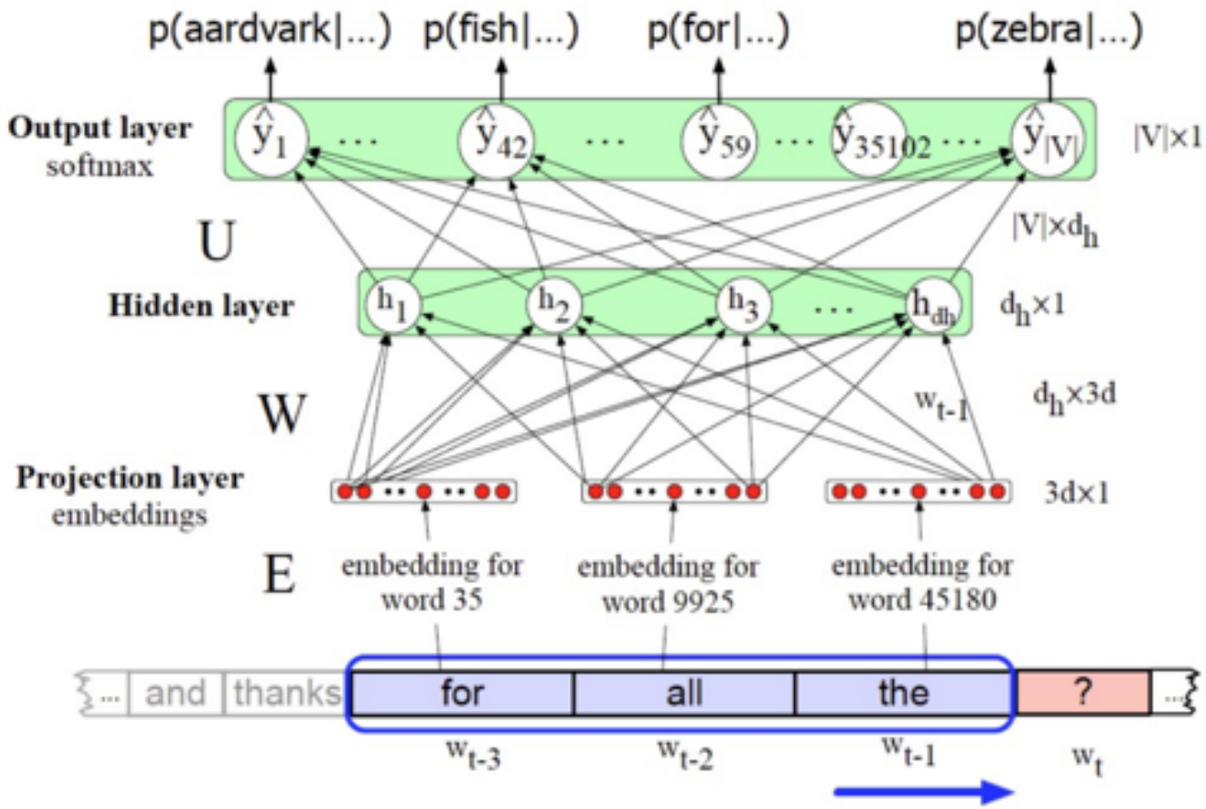
Training data:
We've seen: granuloma within the **rigth** lung
Never seen: granuloma within the **left** lung
Test data: granuloma within the left ___

N-gram LM can't predict "lung"!
Neural LM can use **similarity of** "**left**" and "**rigth**"
Neural embeddings to generalize and predict “lung” after left

Example inspired by https://web.stanford.edu/~jurafsky/slp3/slides/7_NN_Apr_28_2021.pdf

ihealth | Millennium Institute for Intelligent Healthcare Engineering

How do Neural LM learn contextual representations ?

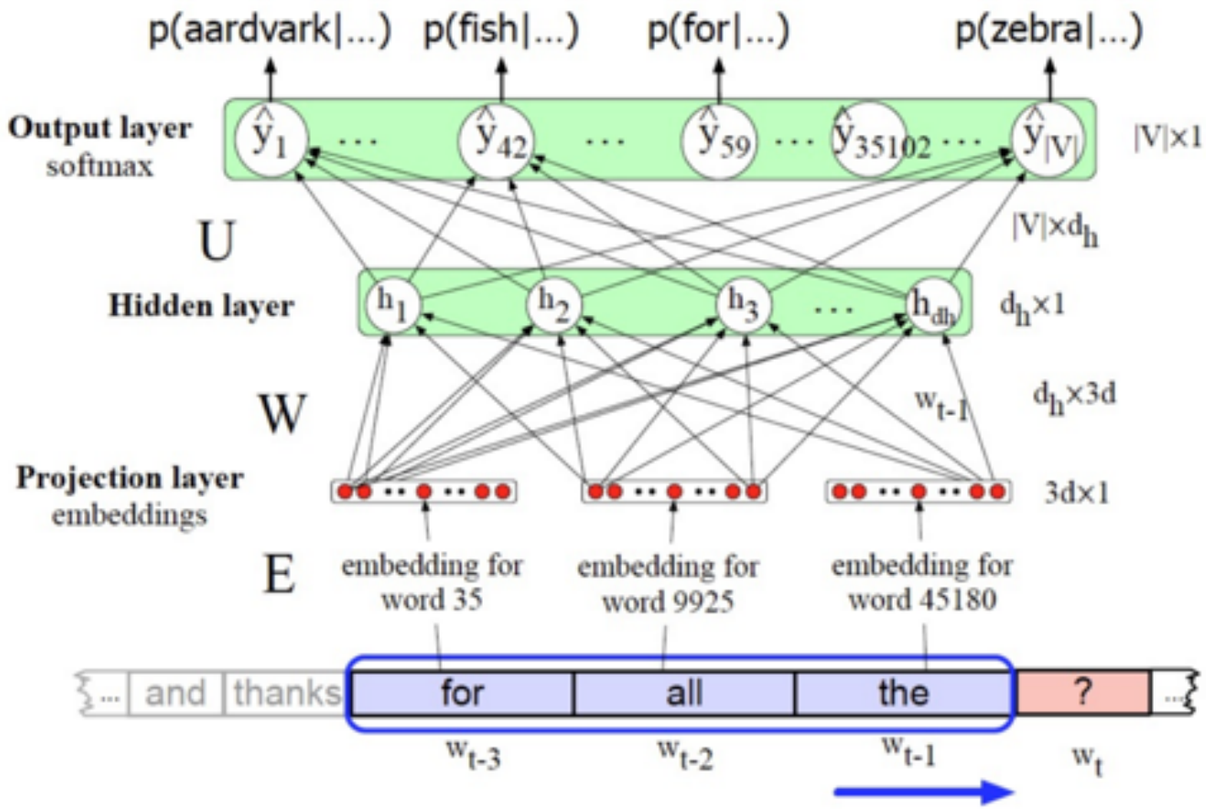


We are learning vectors to represent words and phrases

Image from https://web.stanford.edu/~jurafsky/slp3/slides/7_NN_Apr_28_2021.pdf

ihealth | Millennium Institute for Intelligent Healthcare Engineering

How do Neural LM learn contextual representations ?



If two words have similar vectors, that means they might have similar meaning

Image from https://web.stanford.edu/~jurafsky/slp3/slides/7_NN_Apr_28_2021.pdf

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Word vectors

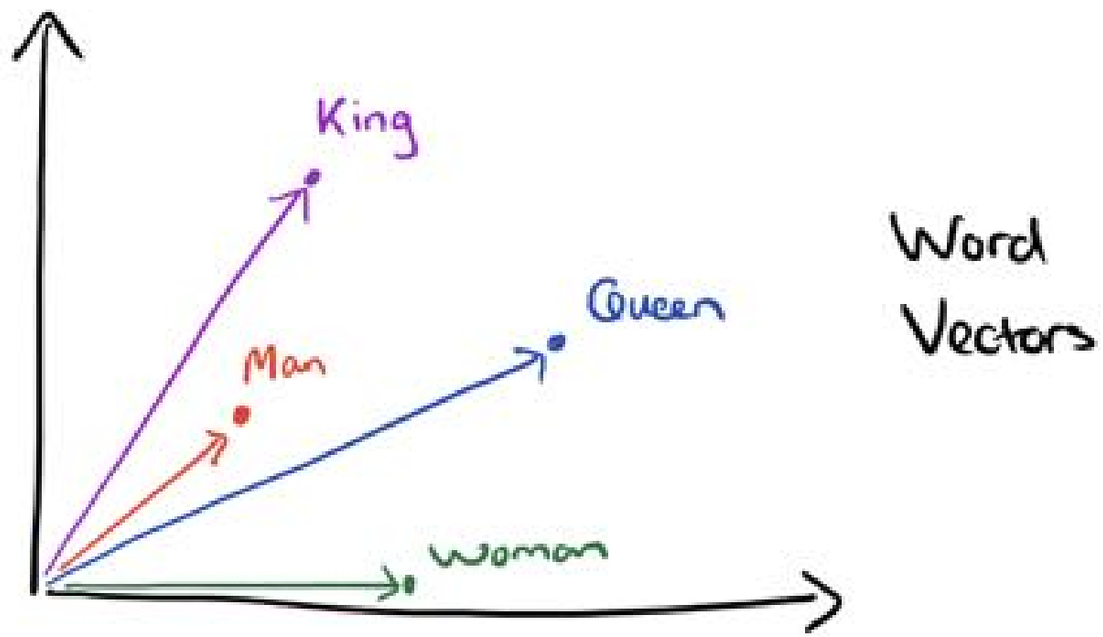
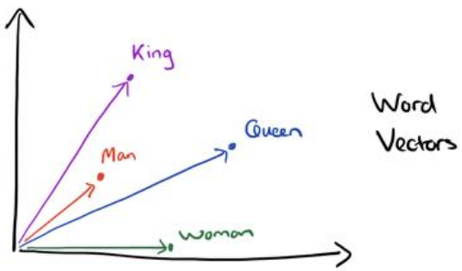


Image from https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Word vectors: composition



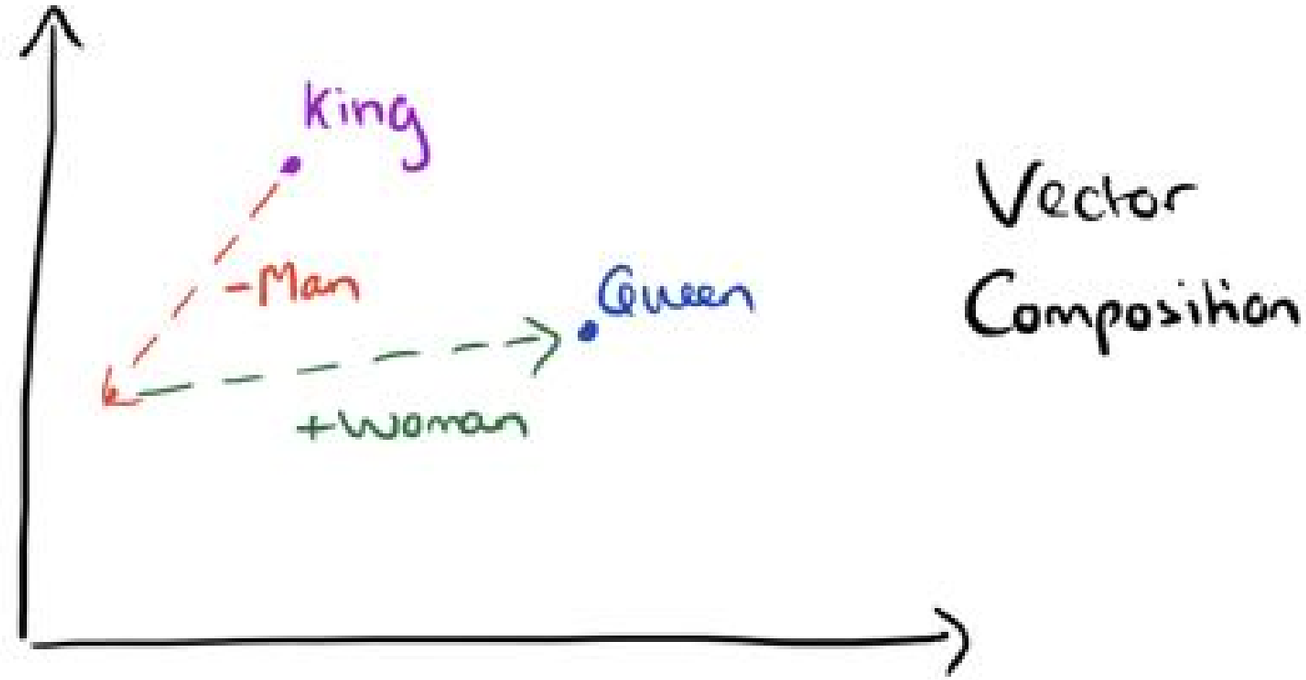


Image from https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Word/Sentence vectors

Johnson, M., Schuster, M., Le, Q.V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G. and Hughes, M., 2017. **Google's multilingual neural machine translation system: Enabling zero-shot translation**. *Transactions of the Association for Computational Linguistics*, *5*, pp.339-351.

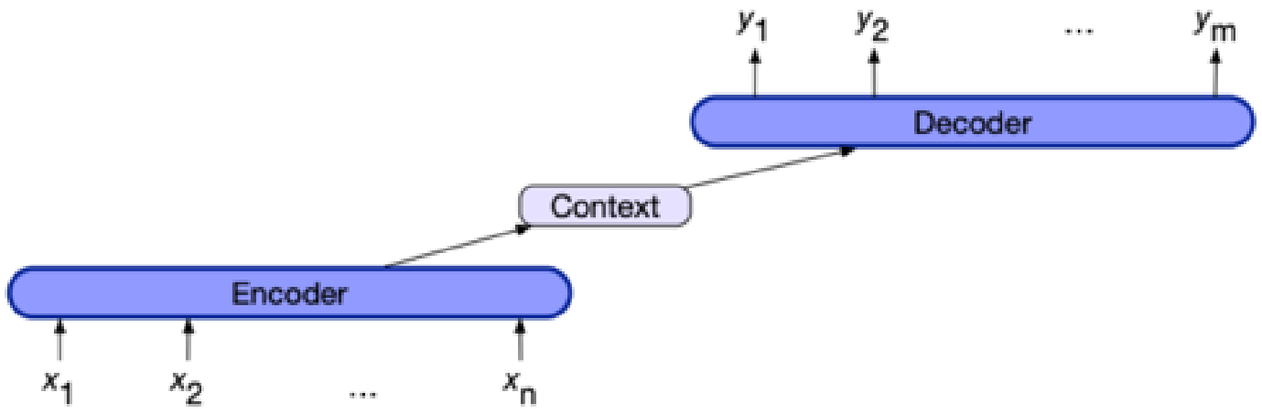
https://aclanthology.org/Q17-1024.pdf

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Neural Architectures for NLG with the Encoder/Decoder model

The encoder-decoder model

Encoder-decoder networks, or **sequence-to-sequence** networks, are models capable of generating contextually appropriate, arbitrary length, output sequences

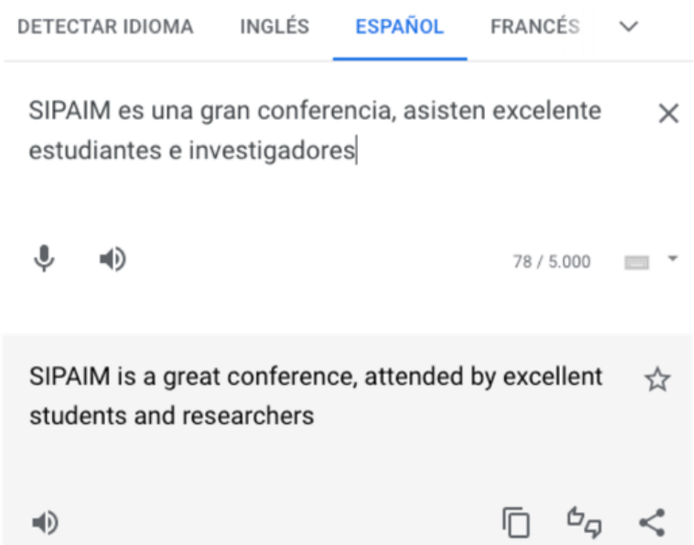


The encoder-decoder architecture. The **context** is a function of the hidden representations of the input, and may be used by the decoder in a variety of ways.

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Applications of E-D model

Machine Translation



Summarization

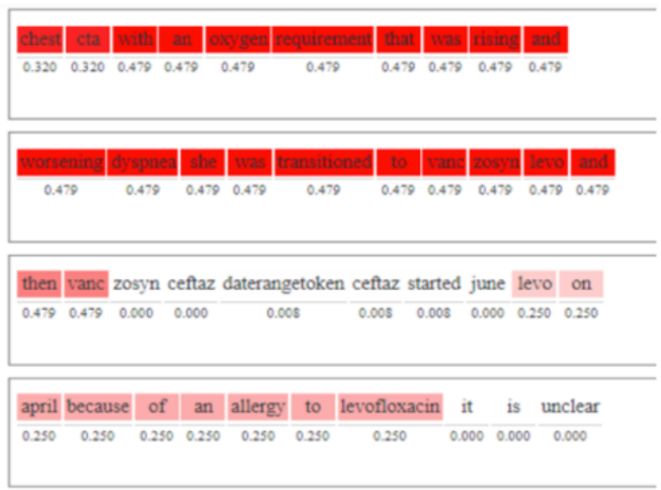


Figure 4: An excerpt of heat-mapping on Neat-Vision tool with transformed attention values to highlight the importance of sentence with red color.

Question Answering

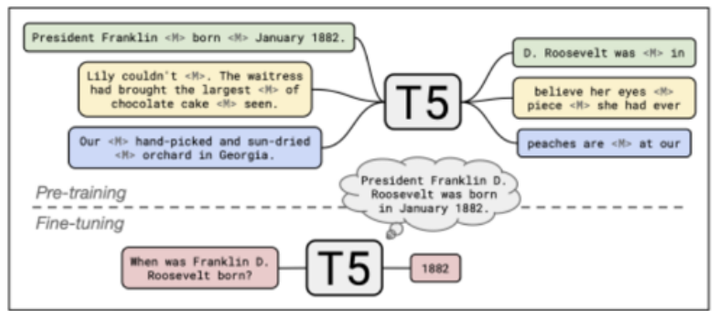


Figure 23.16 The T5 system is an encoder-decoder architecture. In pretraining, it learns to fill in masked spans of task (marked by <M>) by generating the missing spans (separated by <M>) in the decoder. It is then fine-tuned on QA datasets, given the question, without adding any additional context or passages. Figure from Roberts et al. (2020).

ihealth | Millennium Institute for Intelligent Healthcare Engineering

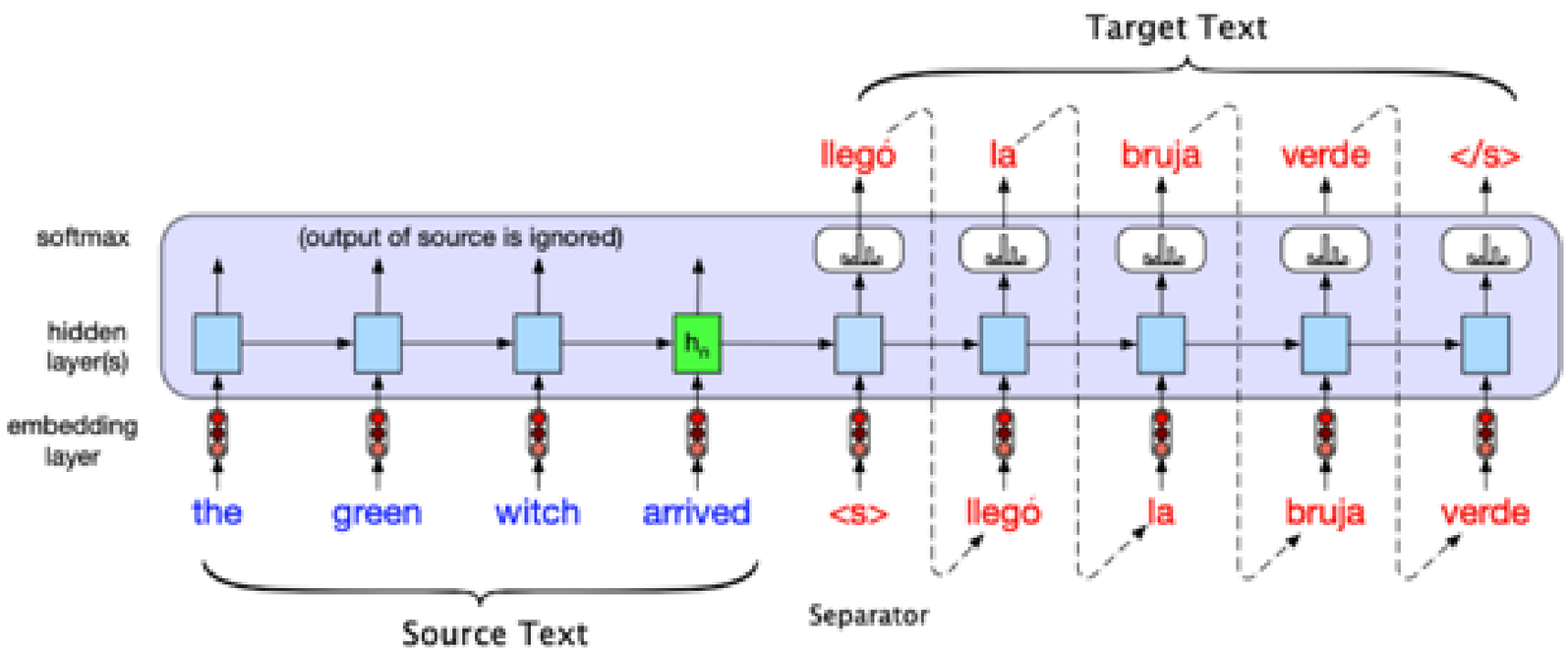
Recurrent architectures

ihealth | Millennium Institute for Intelligent Healthcare Engineering

The encoder-decoder with RNNs

The general E-D model can be implemented with different NN architectures.

An initial version can be implemented with RNNs



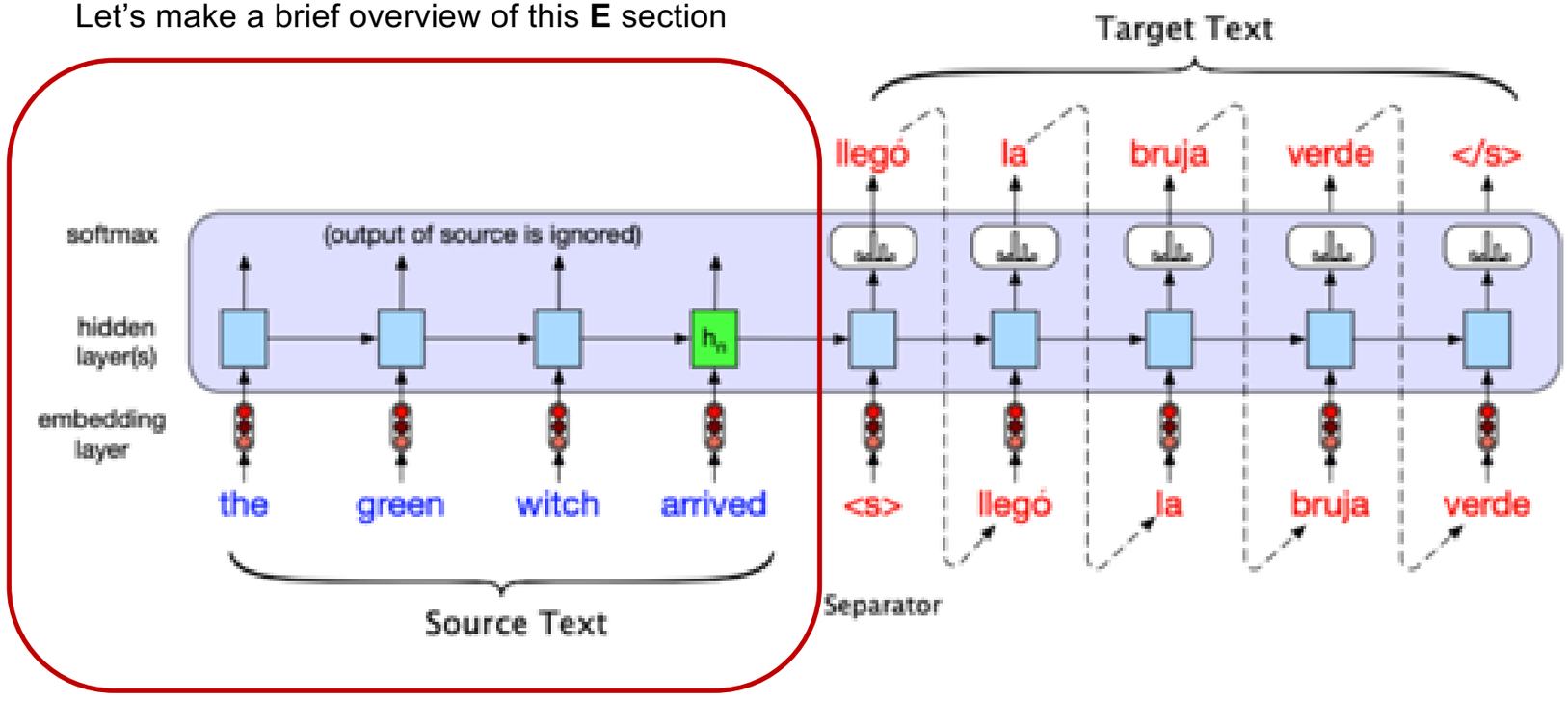
ihealth | Millennium Institute for Intelligent Healthcare Engineering

The encoder-decoder with RNNs

The general E-D model can be implemented with different NN architectures.

An initial version can be implemented with RNNs

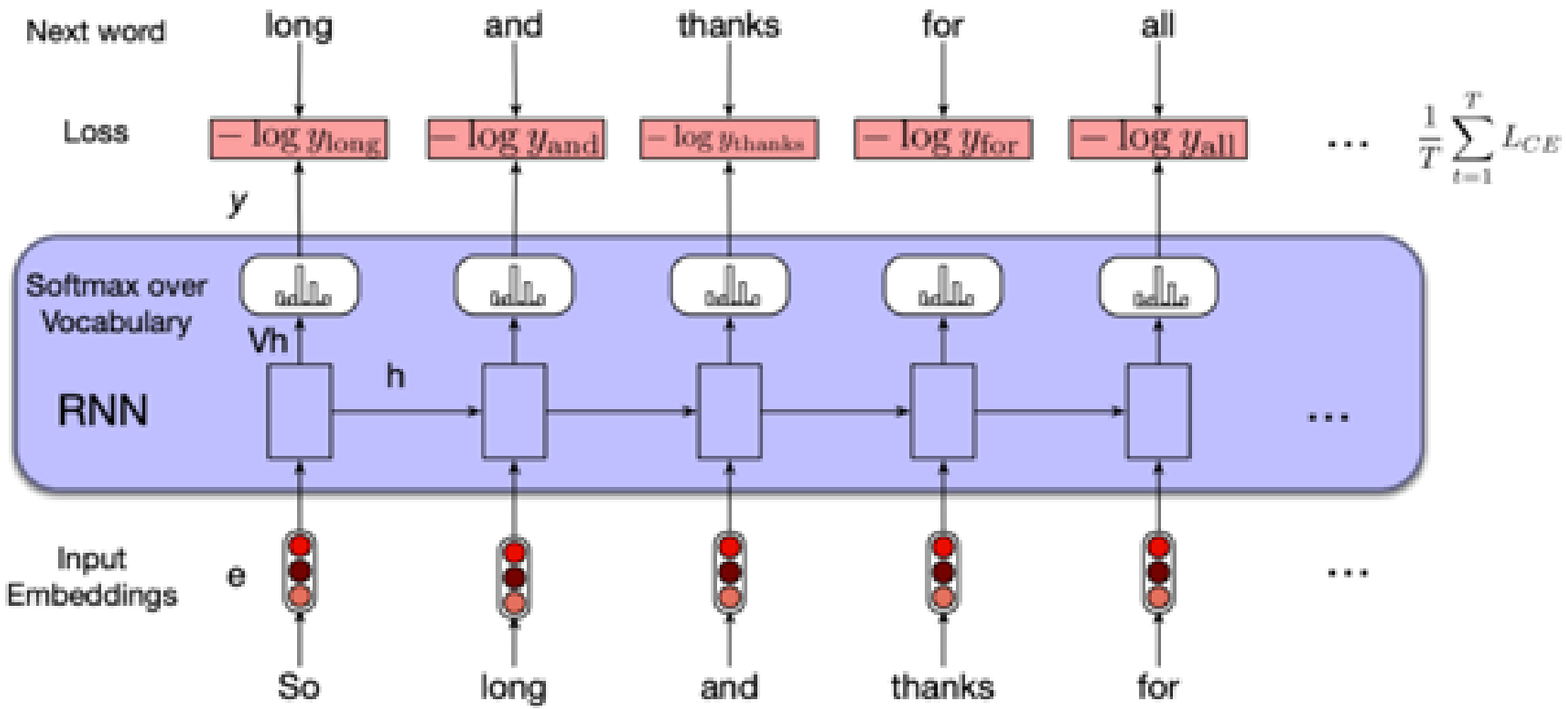
Let's make a brief overview of this **E** section



The RNN Encoder

The RNN Encoder has different components from input to output

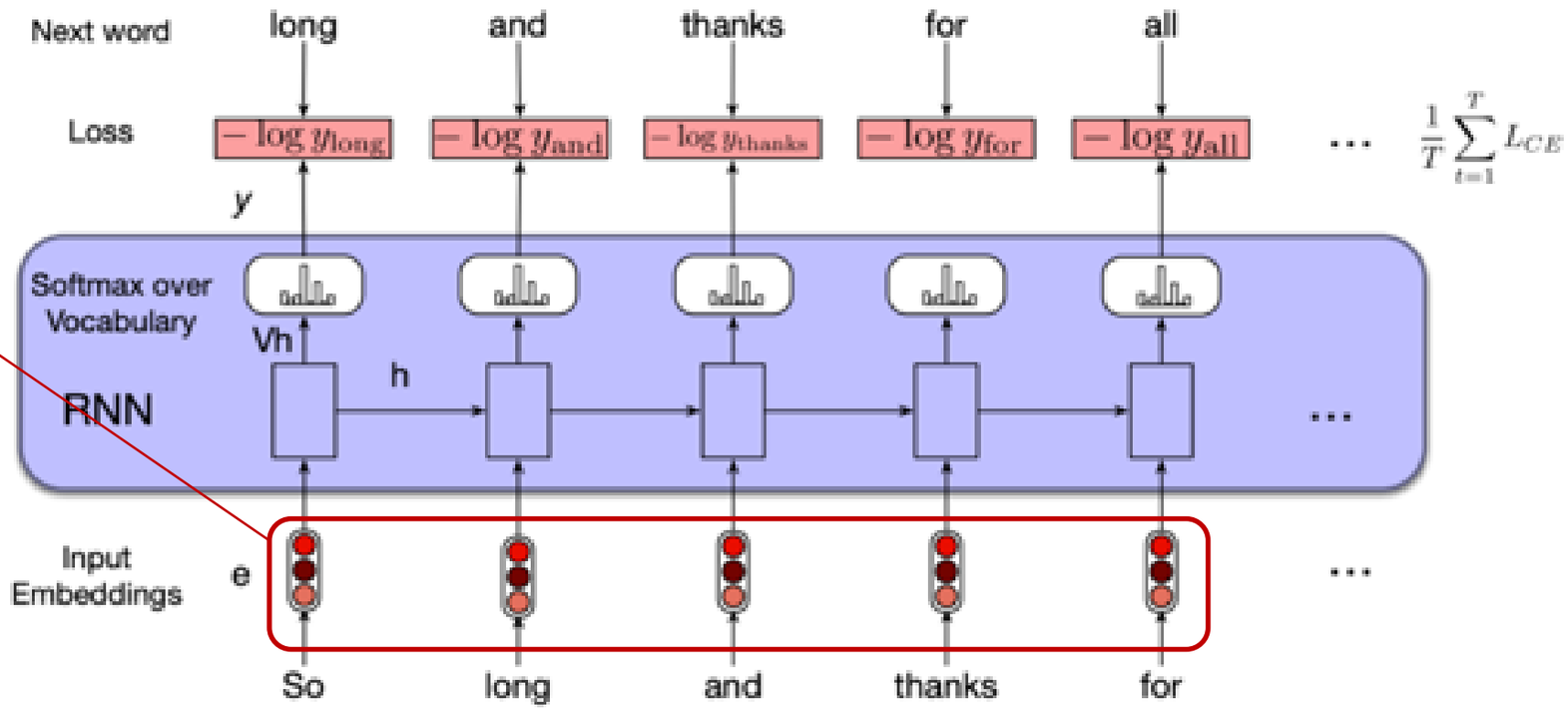
Let's examine them



Training RNNs as language models.

The RNN Encoder

These input word vectors can be trained with the full model or can be pretrained with a word vector model such as word2vec, GloVe or fastext

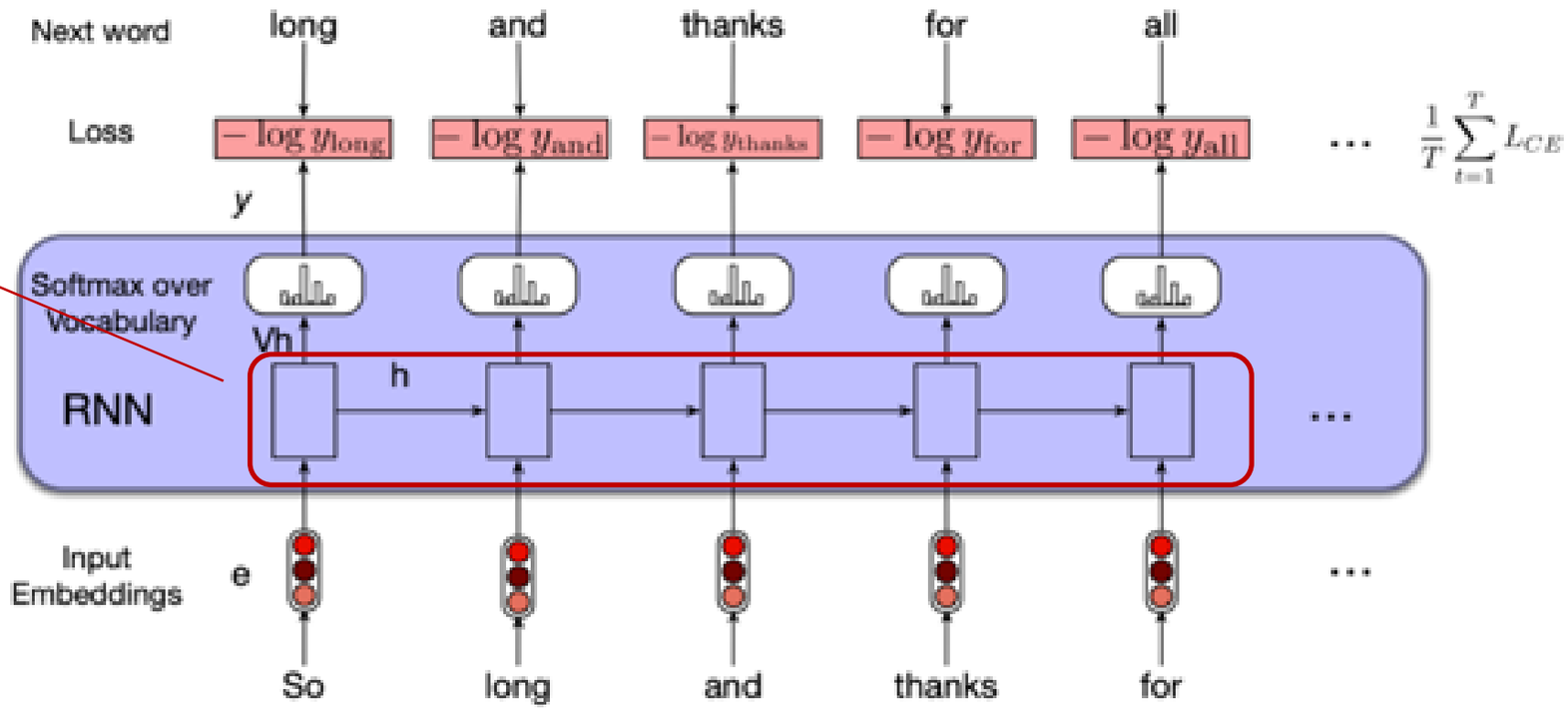


Training RNNs as language models.

The RNN Encoder

These are the RNN cells. There are different types such as the GRU cell or the LSTM cell.

Notice that you need the value of the previous cell (h_{i-1}) to train the value of the next cell (h_i)



Training RNNs as language models.

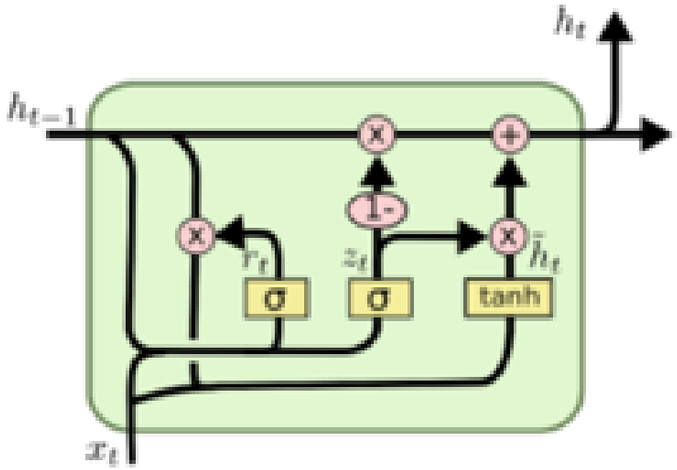
ihealth | Millennium Institute for Intelligent Healthcare Engineering

Detail: The GRU Cell

These cells allow to forget and remember selectively

They also alleviate the vanishing gradient problem

GRU trains a faster than LSTM since it has fewer parameters

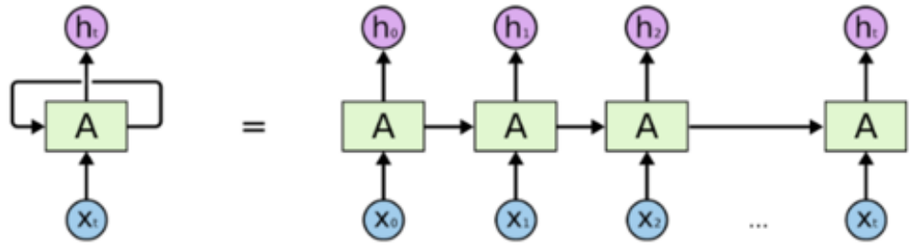


$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

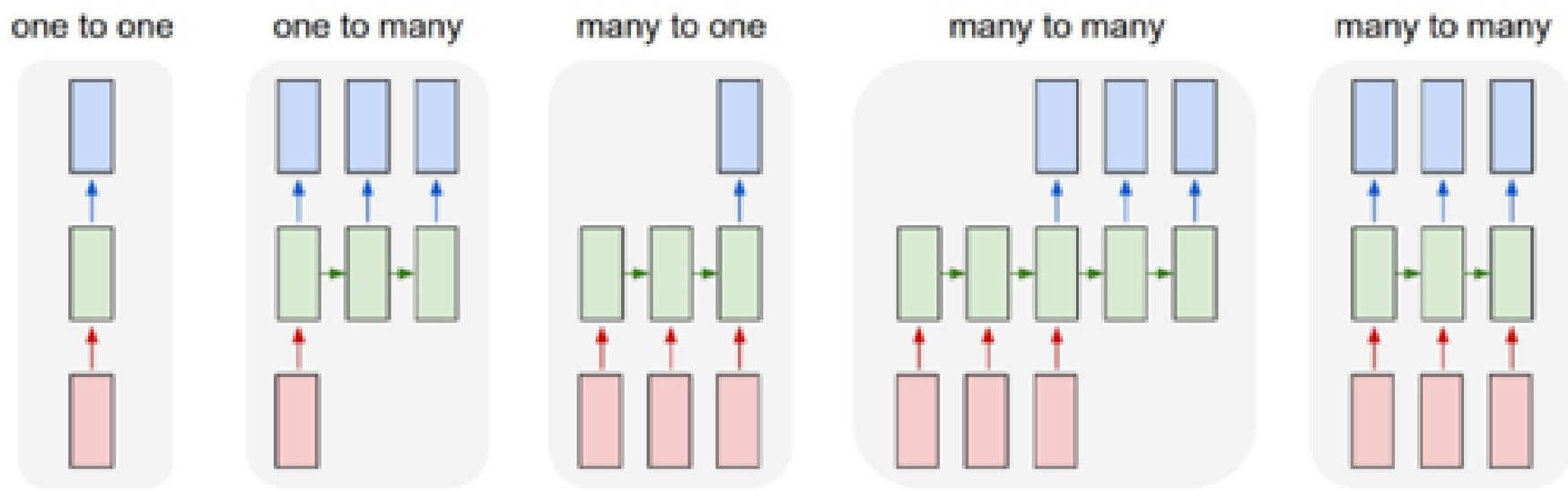
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$



ihealth | Millennium Institute for Intelligent Healthcare Engineering

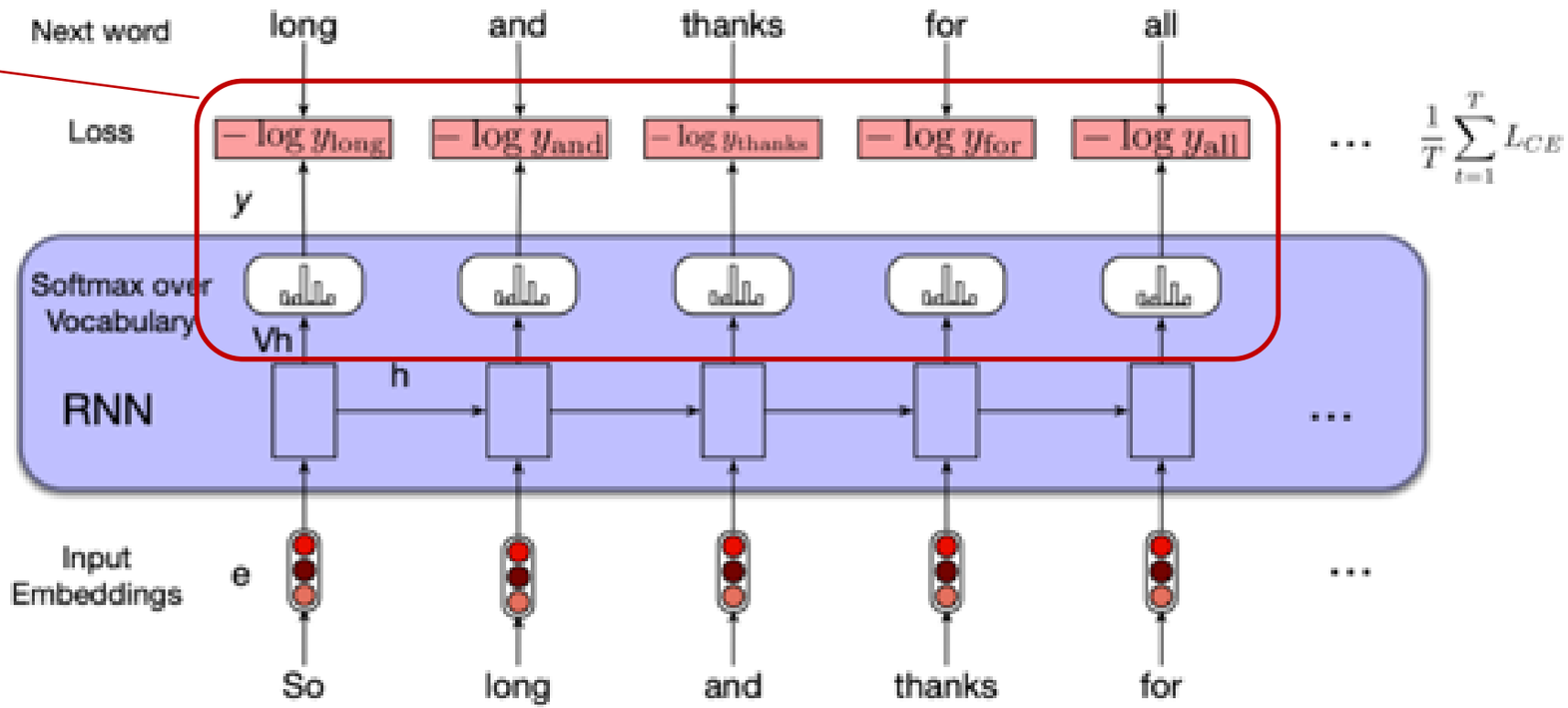
The RNN architecture is very flexible

The RNN architecture is very flexible and allows different types of inputs and outputs, always in sequence



The RNN Encoder

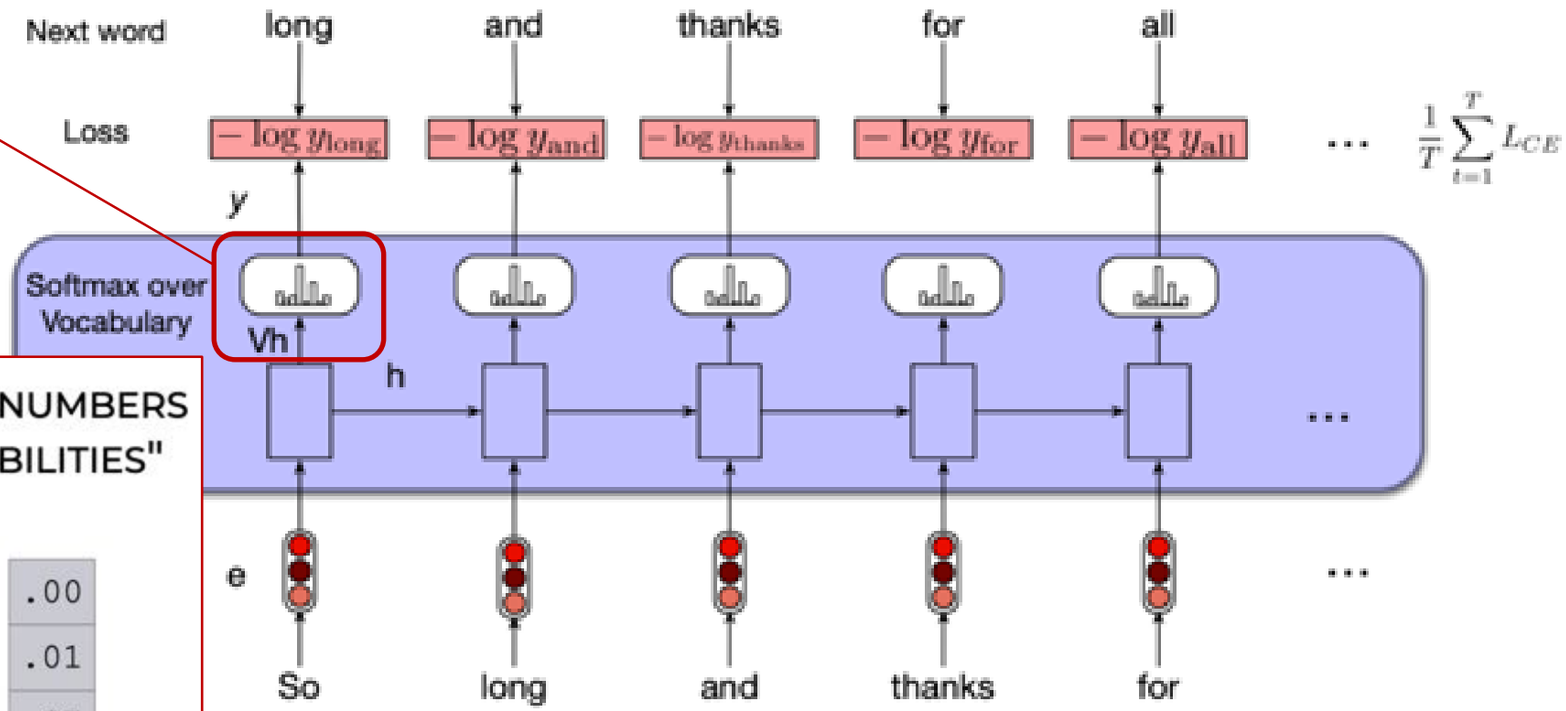
On each step we try to predict the most probable next word given a current sequence



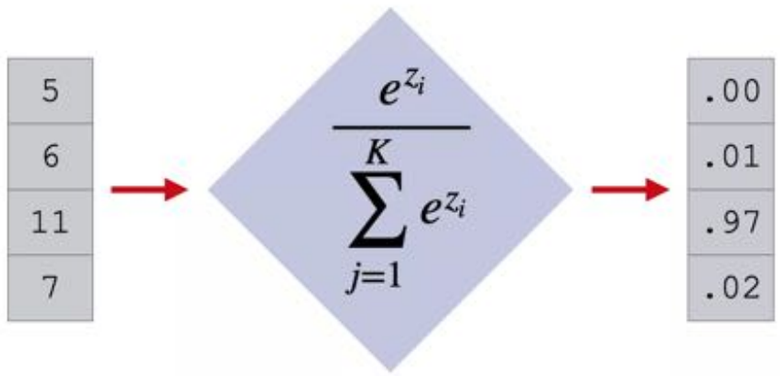
ihealth | Millennium Institute for Intelligent Healthcare Engineering

The RNN Encoder

SOFTMAX gives a probability for each word in the vocabulary



SOFTMAX TRANSFORMS A VECTOR OF NUMBERS INTO A VECTOR OF RELATIVE "PROBABILITIES"



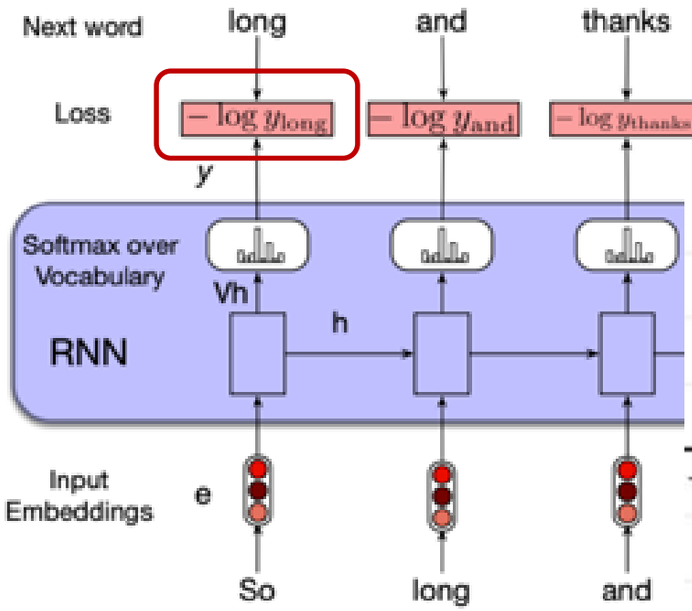
https://www.sharpsightlabs.com/blog/numpy-softmax/

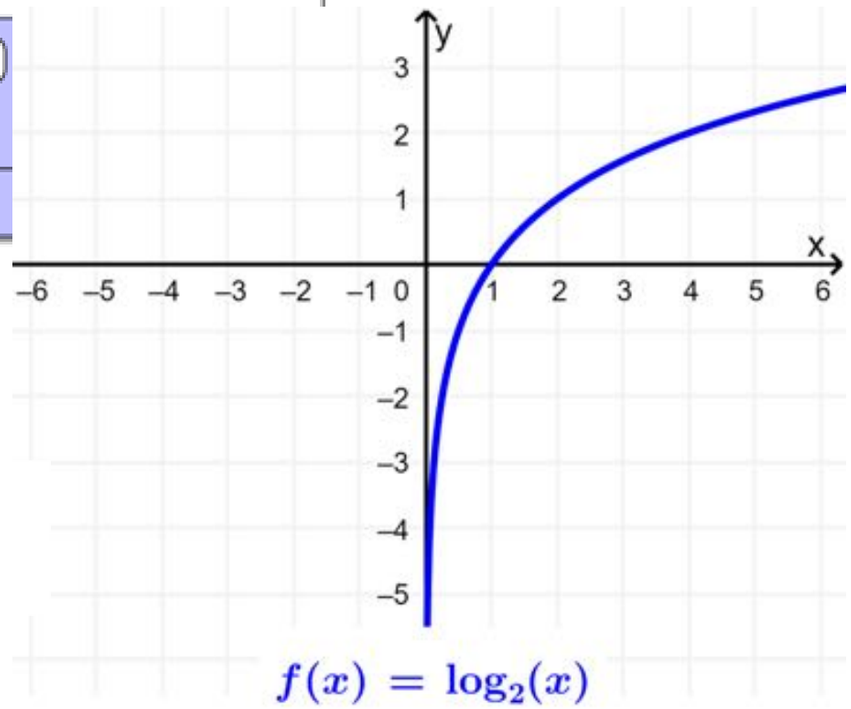
ihealth | Millennium Institute for Intelligent Healthcare Engineering

Cross Entropy Loss

We measure the error with a Cross-Entropy Loss (L_{CE})

In LM, it corresponds to the negative of the log() of the probability of the <u>correct next word</u>





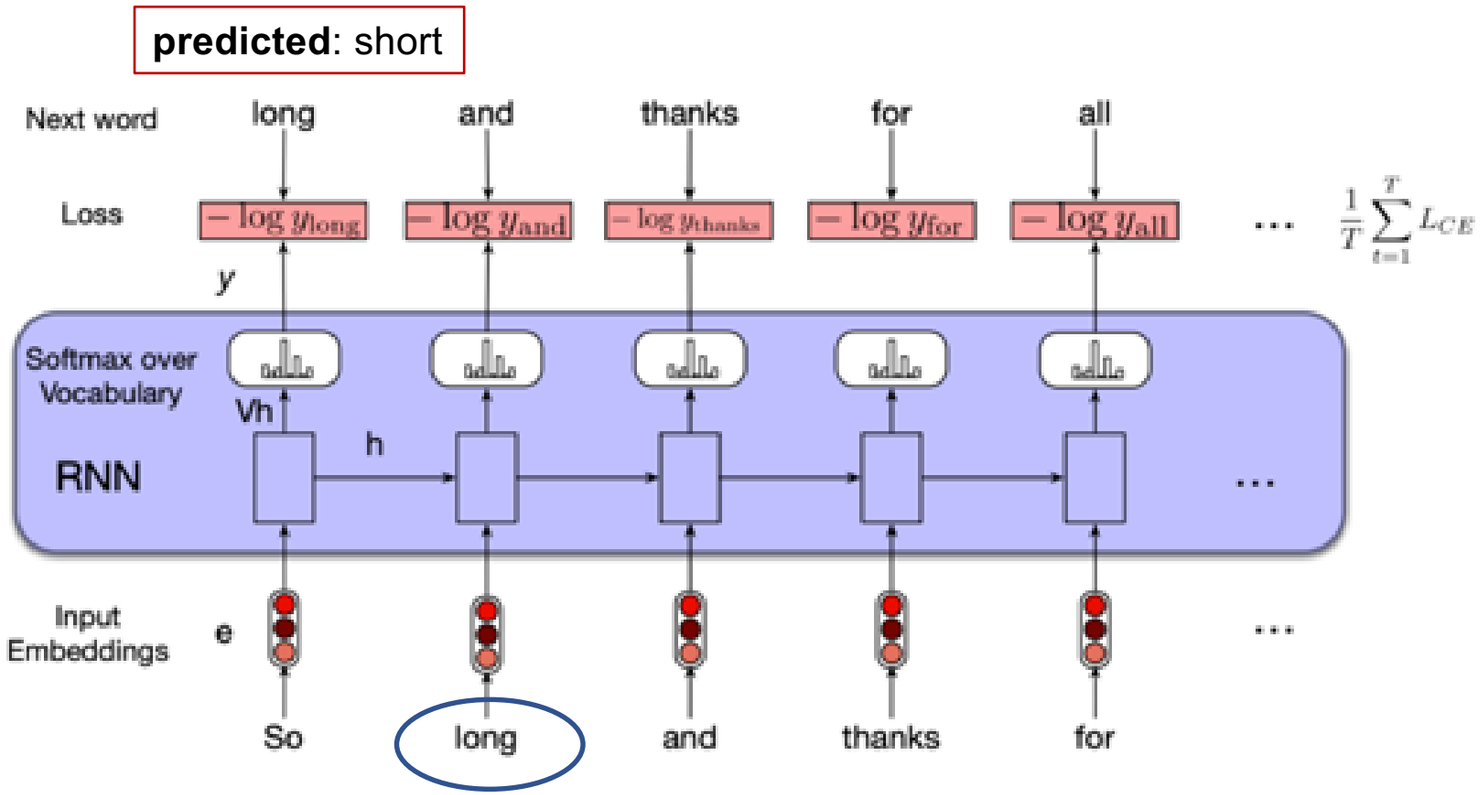
$$L_{CE}(\hat{\mathbf{y}}_t, \mathbf{y}_t) = -\log \hat{\mathbf{y}}_t[w_{t+1}]$$

$$f(x) = \log_2(x)$$

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Teacher forcing

Since we try to predict the next word with the correct previous word (not the wrongly predicted) we call this a TEACHER FORCING approach



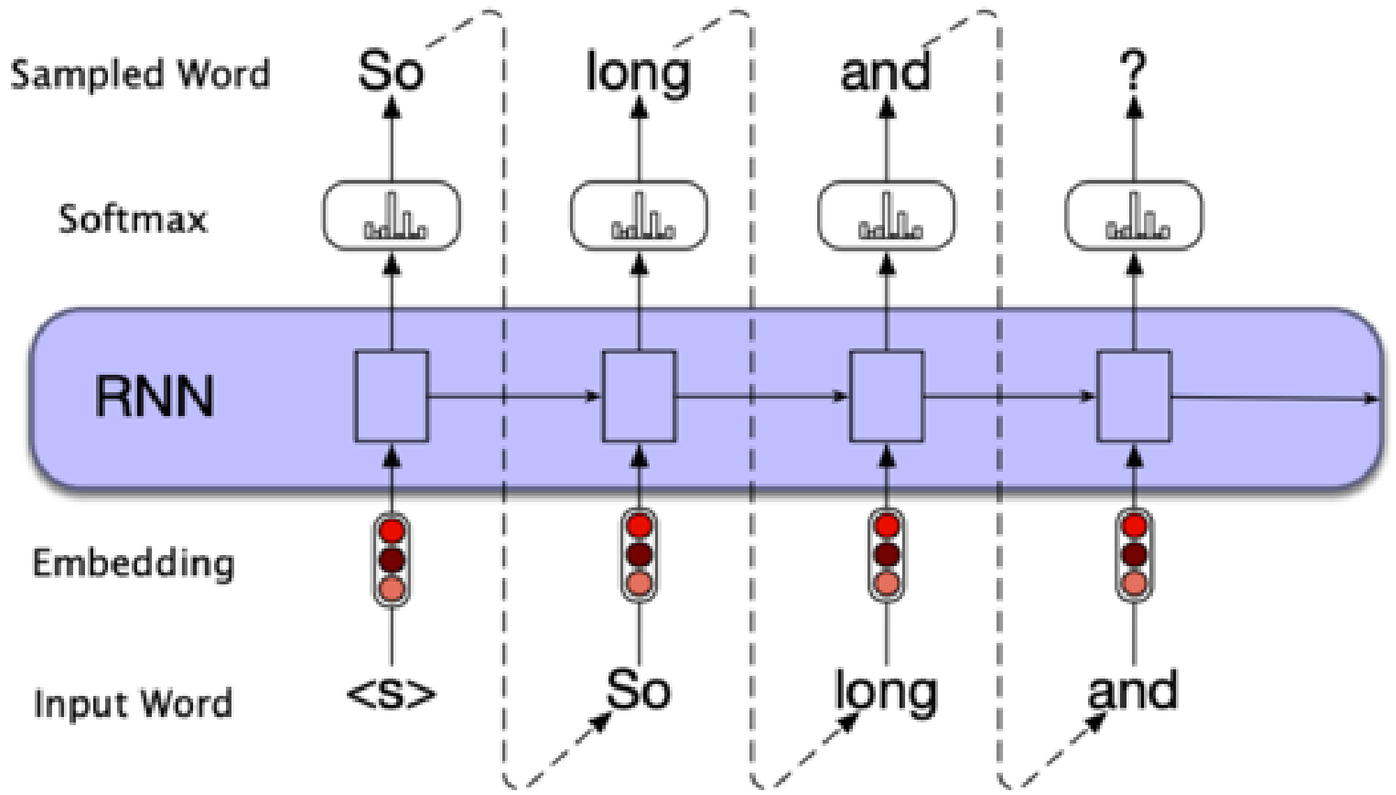
We input “long” anyways

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Back to the Decoder (generator)

Once the RNN is trained, we can use it to generate text by using the output of each step as input to the next one.

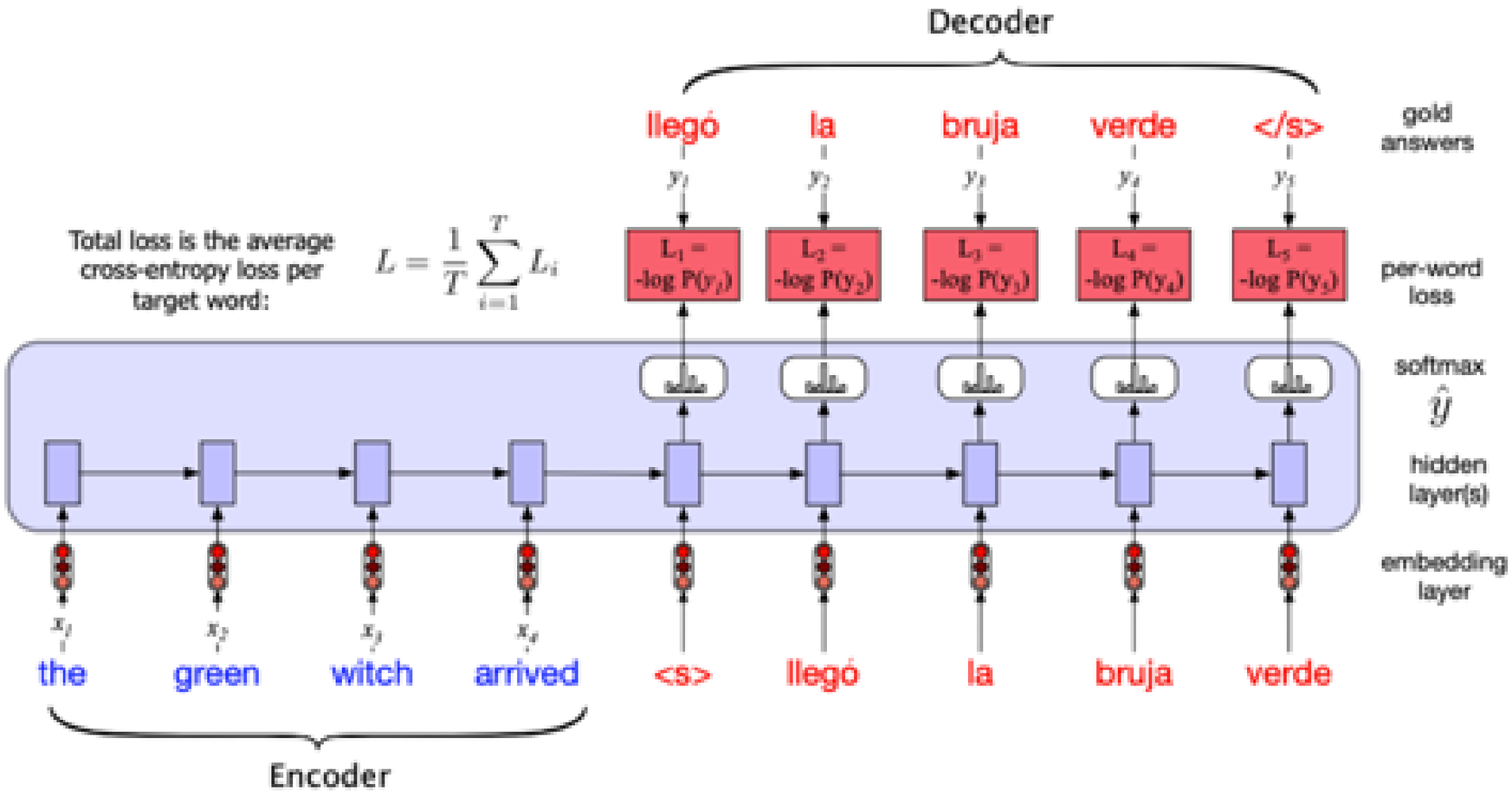
This is called AUTOREGRESSIVE generation



Autoregressive generation with an RNN-based neural language model.

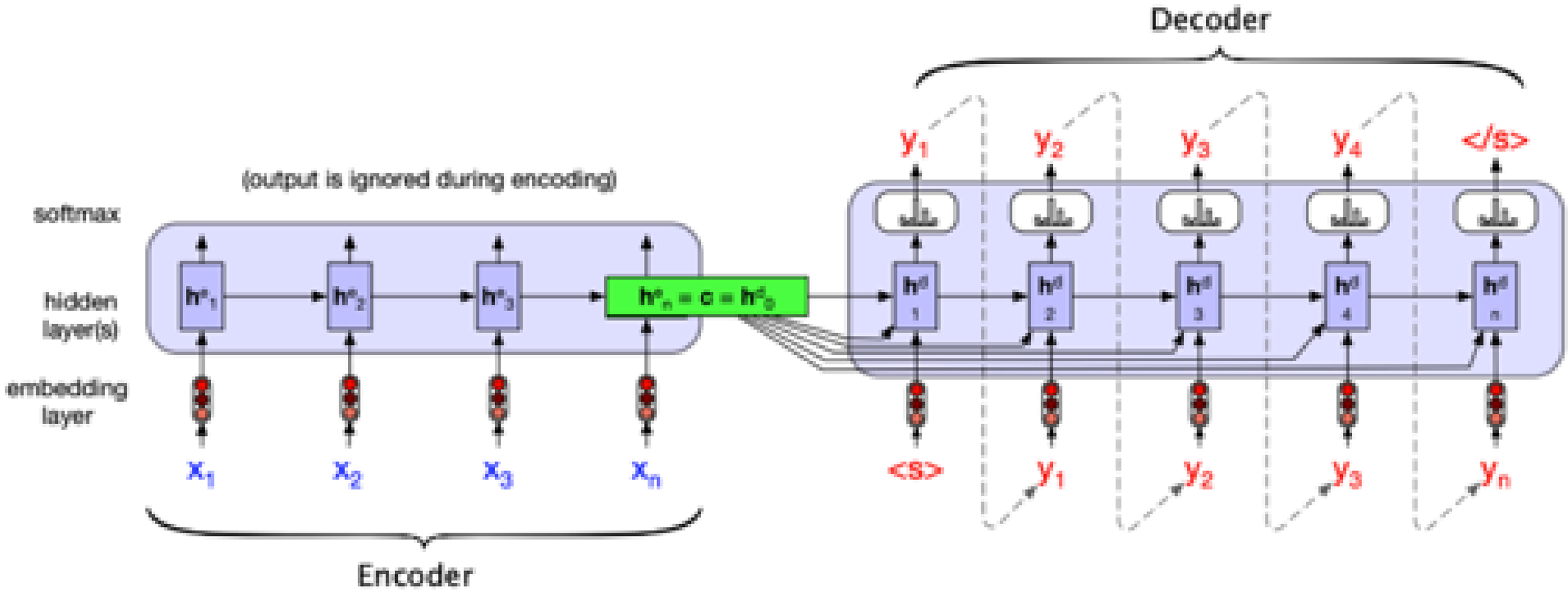
ihealth | Millennium Institute for Intelligent Healthcare Engineering

Training of the full RNN E-D model



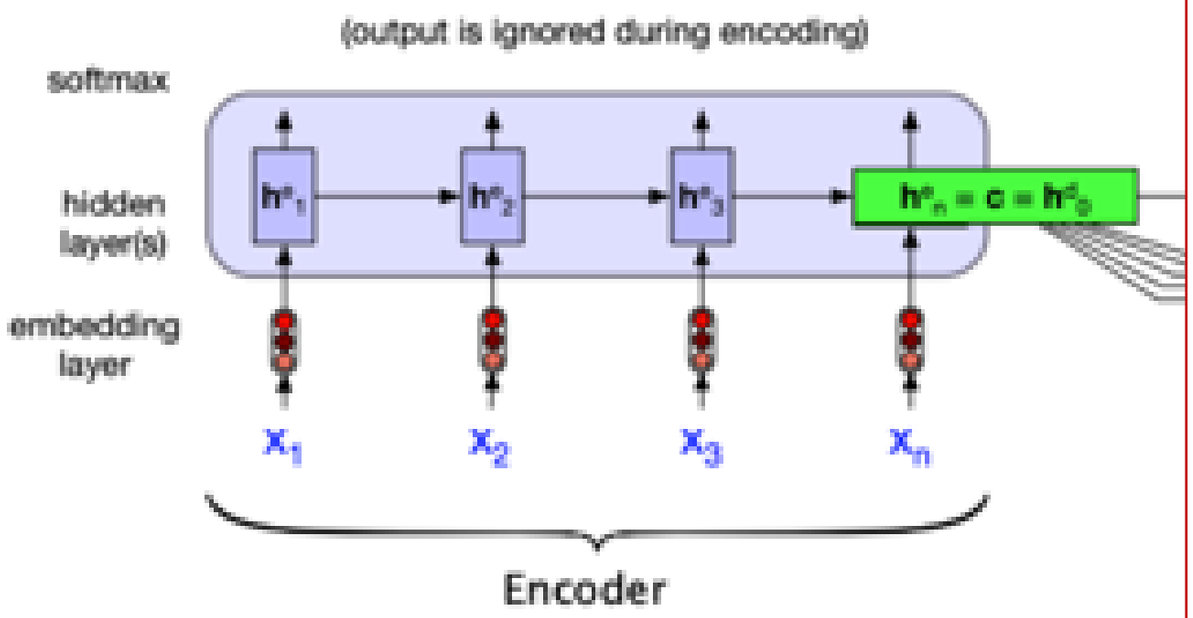
ihealth | Millennium Institute for Intelligent Healthcare Engineering

The full E-D RNN Machine Translation model



ihealth | Millennium Institute for Intelligent Healthcare Engineering

The full E-D RNN Machine Translation model

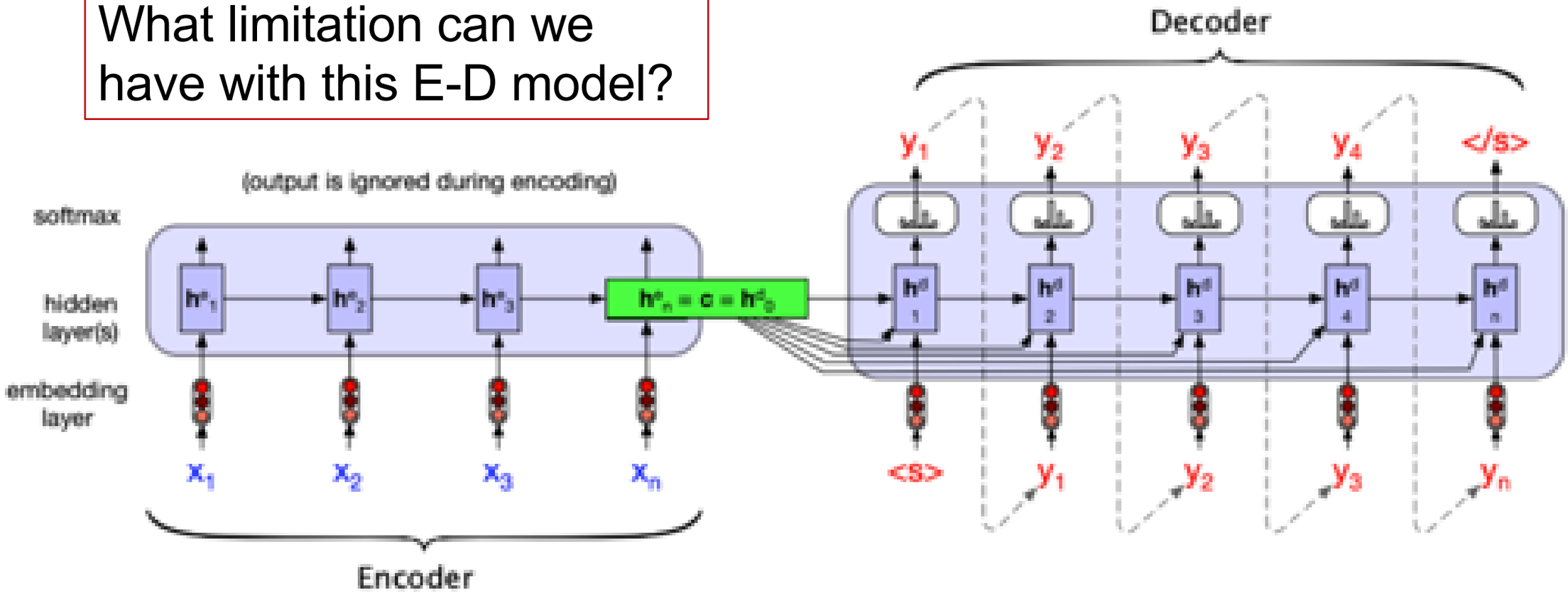


The entire purpose of the **encoder** is to **generate a contextualized representation of the input**. This representation is embodied in the final hidden state of the encoder, $\mathbf{h}^{e}_{n}$. This representation, also called **c** for context, is then passed to the **decoder.**

ihealth | Millennium Institute for Intelligent Healthcare Engineering

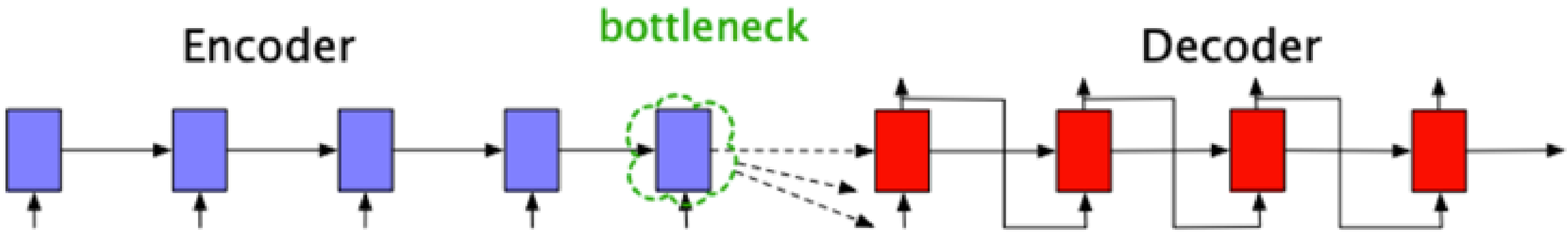
Limitations

What limitation can we have with this E-D model?



ihealth | Millennium Institute for Intelligent Healthcare Engineering

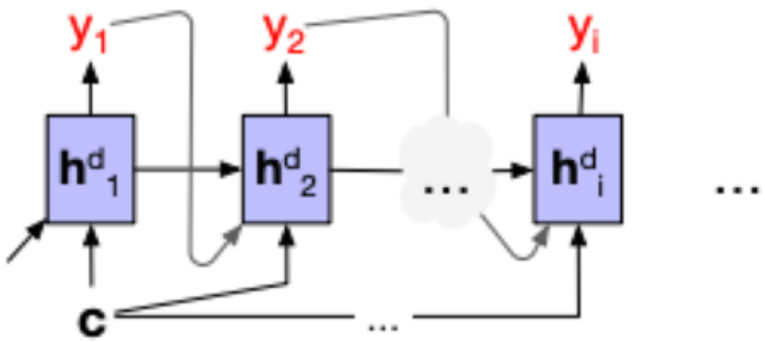
Bottleneck



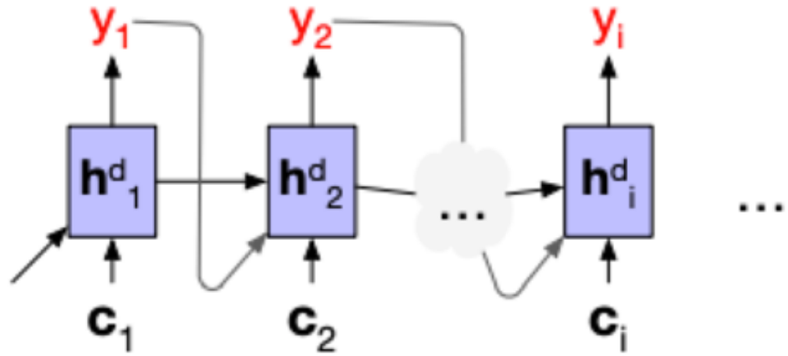
Requiring the context c to be only the encoder's final hidden state forces all the information from the entire source sentence to pass through this representational bottleneck.

ihealth | Millennium Institute for Intelligent Healthcare Engineering

A new context for each decode Hidden state



$$\mathbf{h}_t^d = g(\hat{y}_{t-1}, \mathbf{h}_{t-1}^d, \mathbf{c})$$



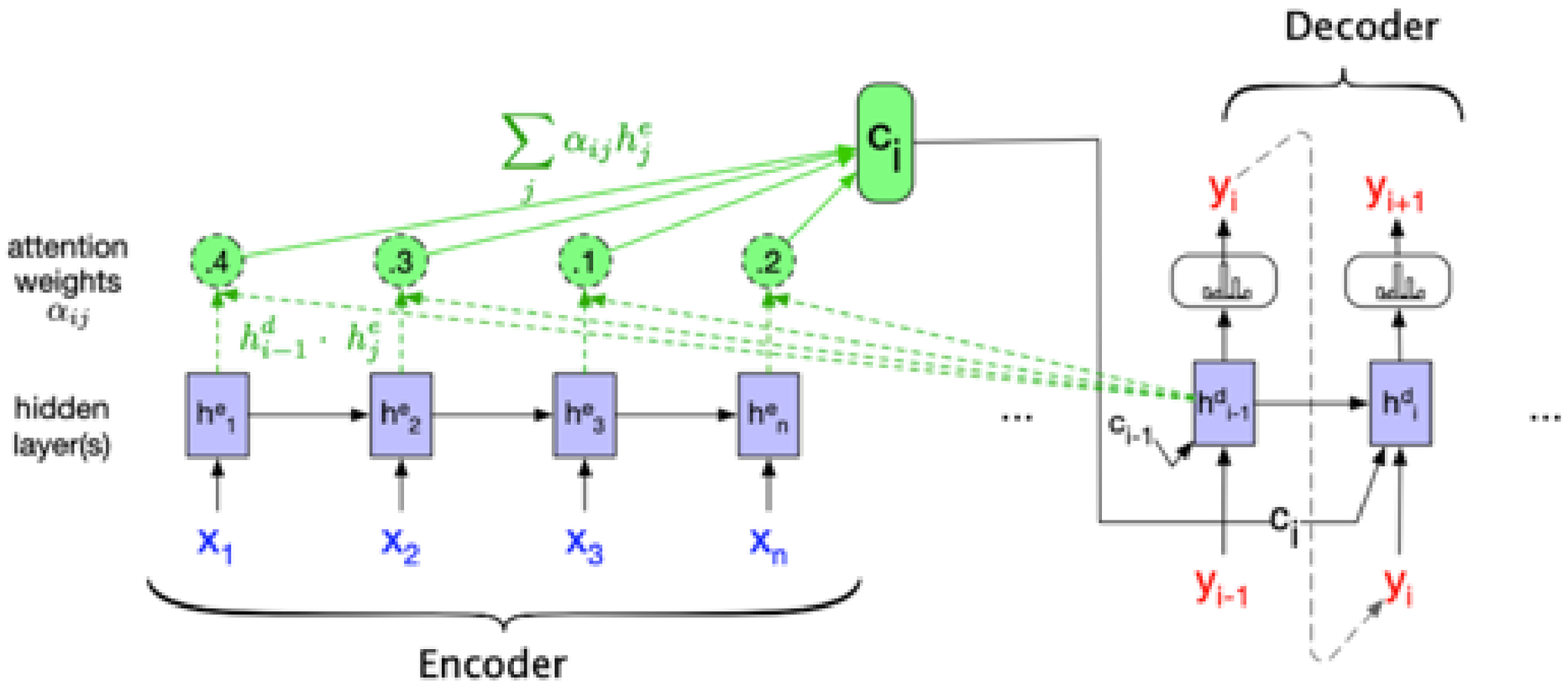
$$\mathbf{h}_i^d = g(\hat{y}_{i-1}, \mathbf{h}_{i-1}^d, \mathbf{c}_i)$$

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Attention

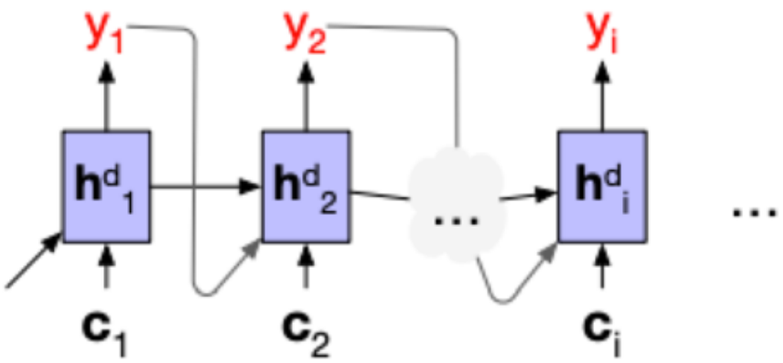
ihealth | Millennium Institute for Intelligent Healthcare Engineering

Encoder Decoder with Attention



ihealth | Millennium Institute for Intelligent Healthcare Engineering

Attention



...

$$\mathbf{h}_i^d = g(\hat{y}_{i-1}, \mathbf{h}_{i-1}^d, \mathbf{c}_i)$$

$$\mathbf{c}_i = \sum_j \alpha_{ij}\, \mathbf{h}_j^e$$

$$\alpha_{ij} = \text{softmax}(score(\mathbf{h}_{i-1}^d, \mathbf{h}_j^e) \;\; \forall j \in e)$$

$$= \frac{\exp(score(\mathbf{h}_{i-1}^d, \mathbf{h}_j^e)}{\sum_k \exp(score(\mathbf{h}_{i-1}^d, \mathbf{h}_k^e))}$$

$$score(\mathbf{h}_{i-1}^d, \mathbf{h}_j^e) = \mathbf{h}_{i-1}^d \cdot \mathbf{h}_j^e$$

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Beam Search

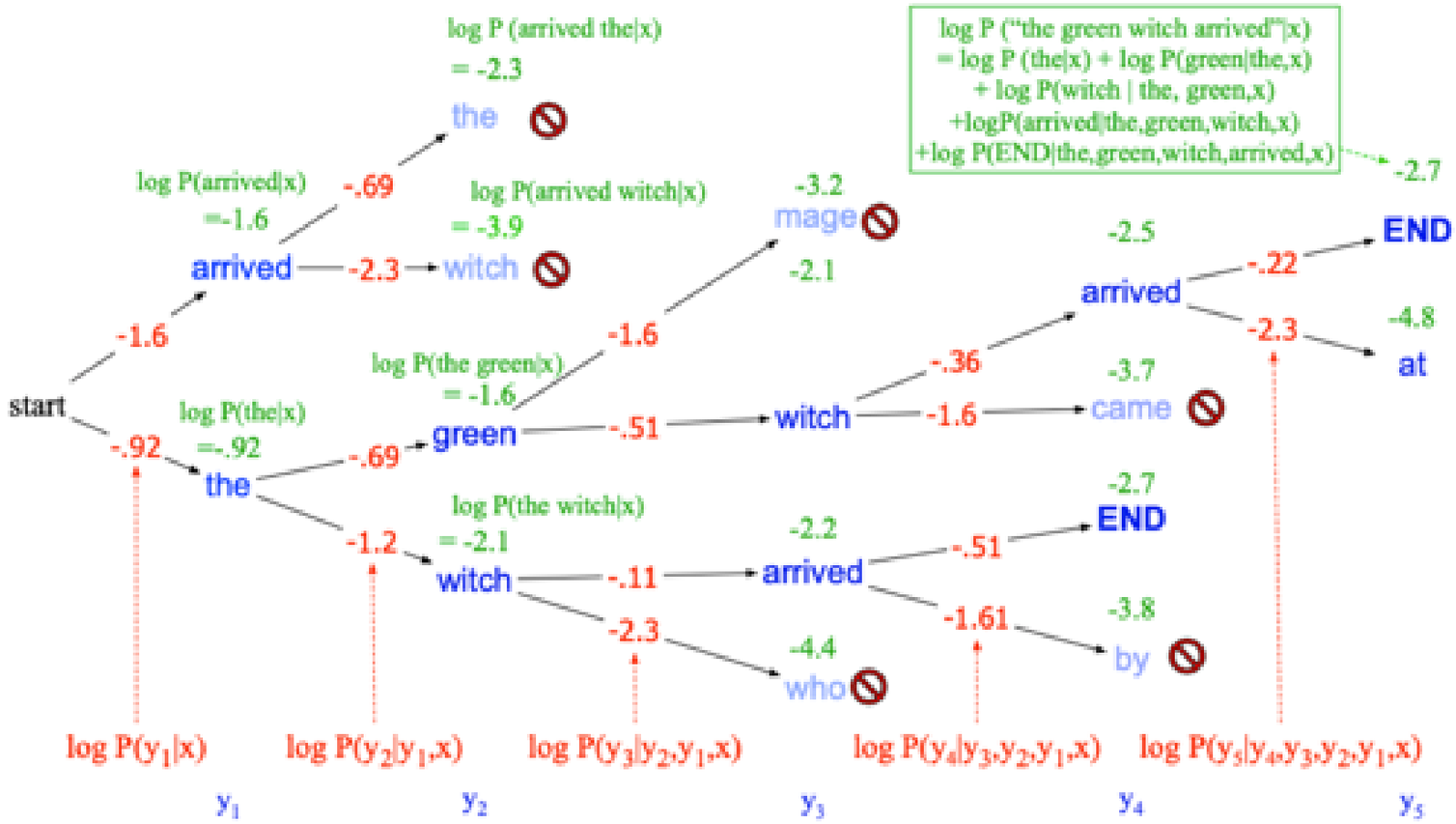
ihealth | Millennium Institute for Intelligent Healthcare Engineering

Beam Search

Until now, we always choose the most probable next word

$$\hat{y}_t = \text{argmax}_{w \in V} P(w|x, y_1 \ldots y_{t-1})$$

Sometimes, we better search for a most probable sequence, even if the next word is not the most probable



ihealth | Millennium Institute for Intelligent Healthcare Engineering

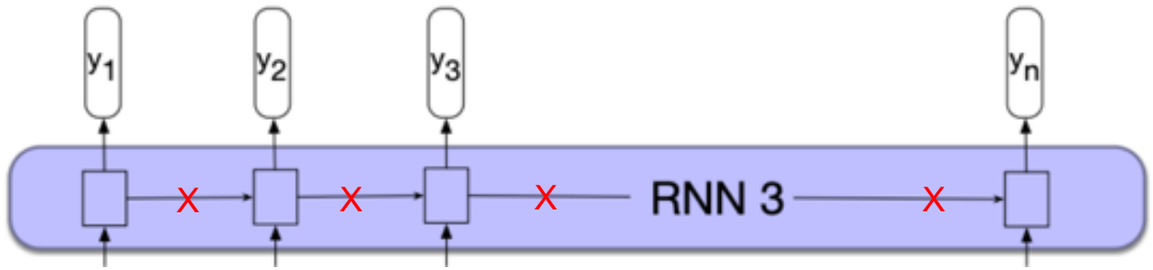
Transformer

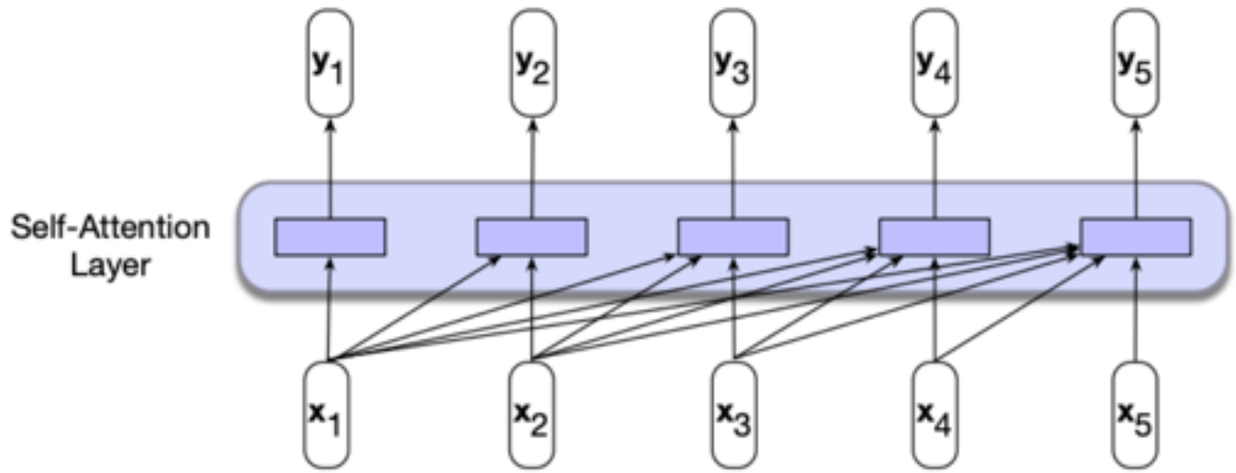
ihealth | Millennium Institute for Intelligent Healthcare Engineering

Brief overview of the Transformer

The inherently sequential nature of recurrent networks makes it hard to do computation in parallel.

These considerations led to the development of transformers**, an approach to sequence processing that eliminates recurrent connections and returns to architectures reminiscent of the fully connected networks**





Self attention

New roles: Key, Query, Value

query

key

value

- As *the current focus of attention* when being compared to all of the other preceding inputs. We'll refer to this role as a **query**.
- In its role as *a preceding input* being compared to the current focus of attention. We'll refer to this role as a **key**.
- And finally, as a **value** used to compute the output for the current focus of attention.

In RNN models

$$\text{score}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \cdot \mathbf{x}_j$$

$$\mathbf{y}_i = \sum_{j \le i} \alpha_{ij} \mathbf{x}_j$$

In transformer models

$$\text{score}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{q}_i \cdot \mathbf{k}_j$$

$$\mathbf{y}_i = \sum_{j \le i} \alpha_{ij} \mathbf{v}_j$$

Self attention

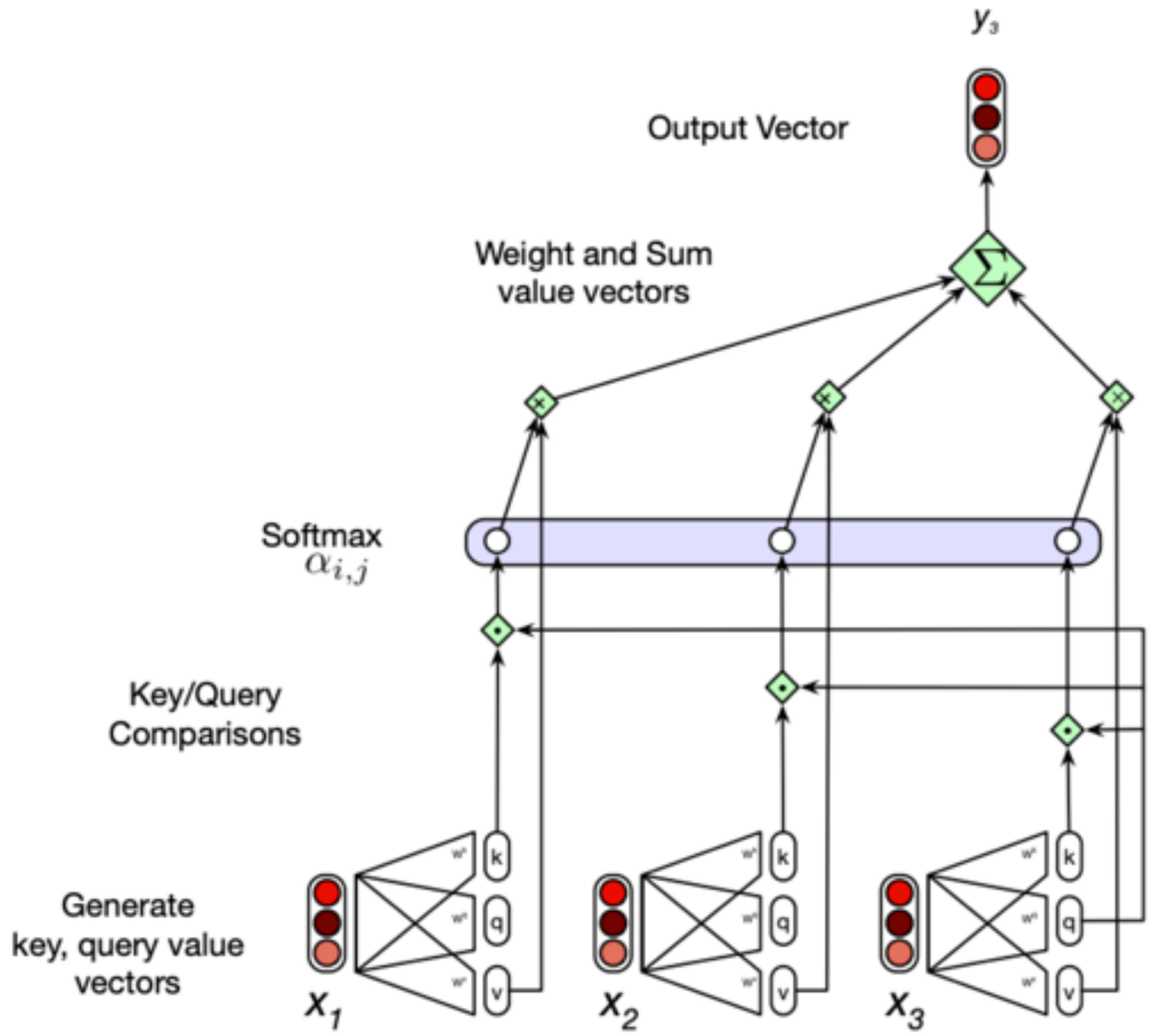
To calculate y_3

$$\mathbf{q}_i = \mathbf{W}^{\mathbf{Q}}\mathbf{x}_i; \;\; \mathbf{k}_i = \mathbf{W}^{\mathbf{K}}\mathbf{x}_i; \;\; \mathbf{v}_i = \mathbf{W}^{\mathbf{V}}\mathbf{x}_i$$

then

$$\text{score}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{q}_i \cdot \mathbf{k}_j$$

$$\mathbf{y}_i = \sum_{j \leq i} \alpha_{ij}\mathbf{v}_j$$

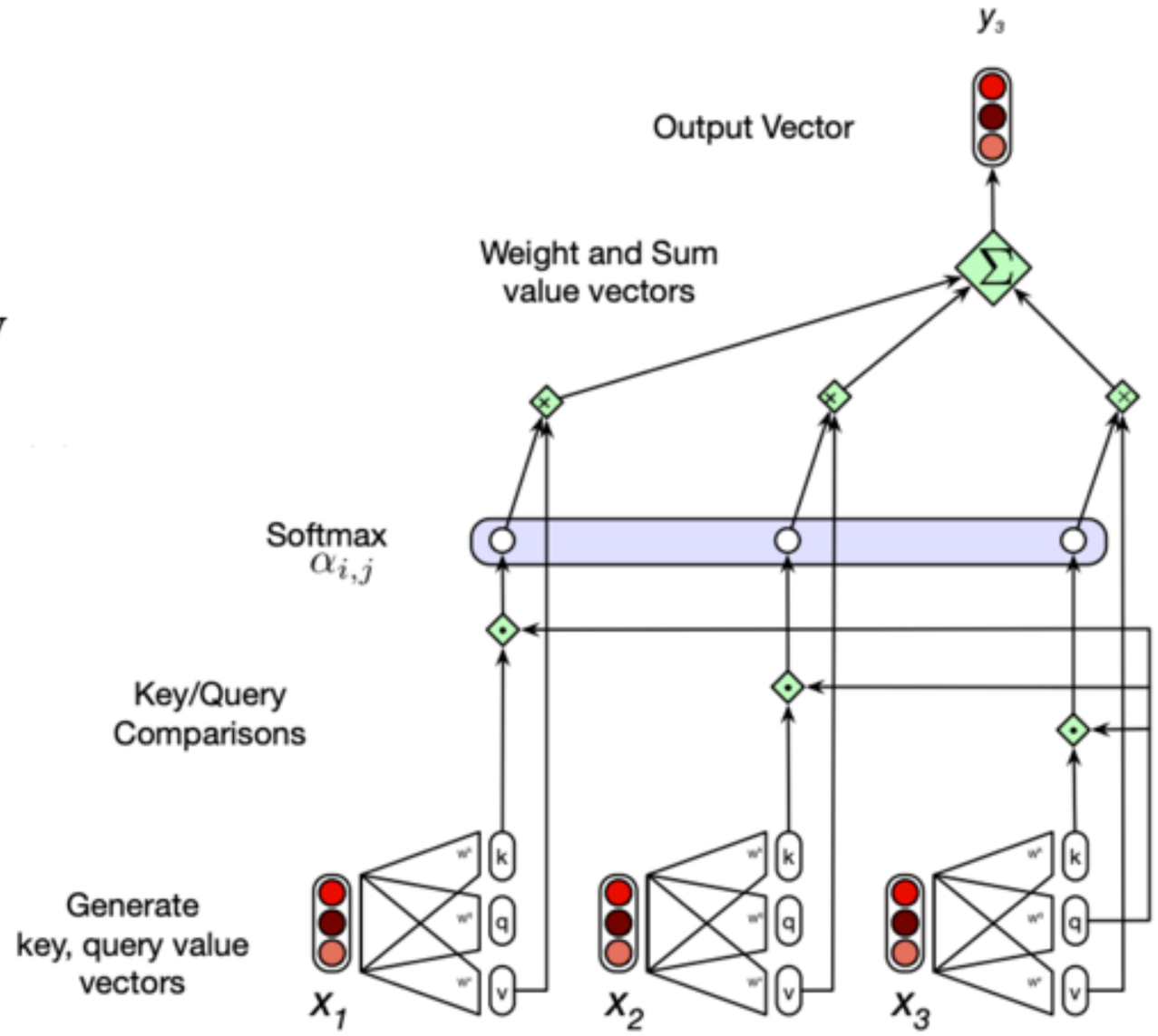


ihealth | Millennium Institute for Intelligent Healthcare Engineering

Self attention

$$SelfAttention(\mathbf{Q},\mathbf{K},\mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^\intercal}{\sqrt{d_k}}\right)\mathbf{V}$$

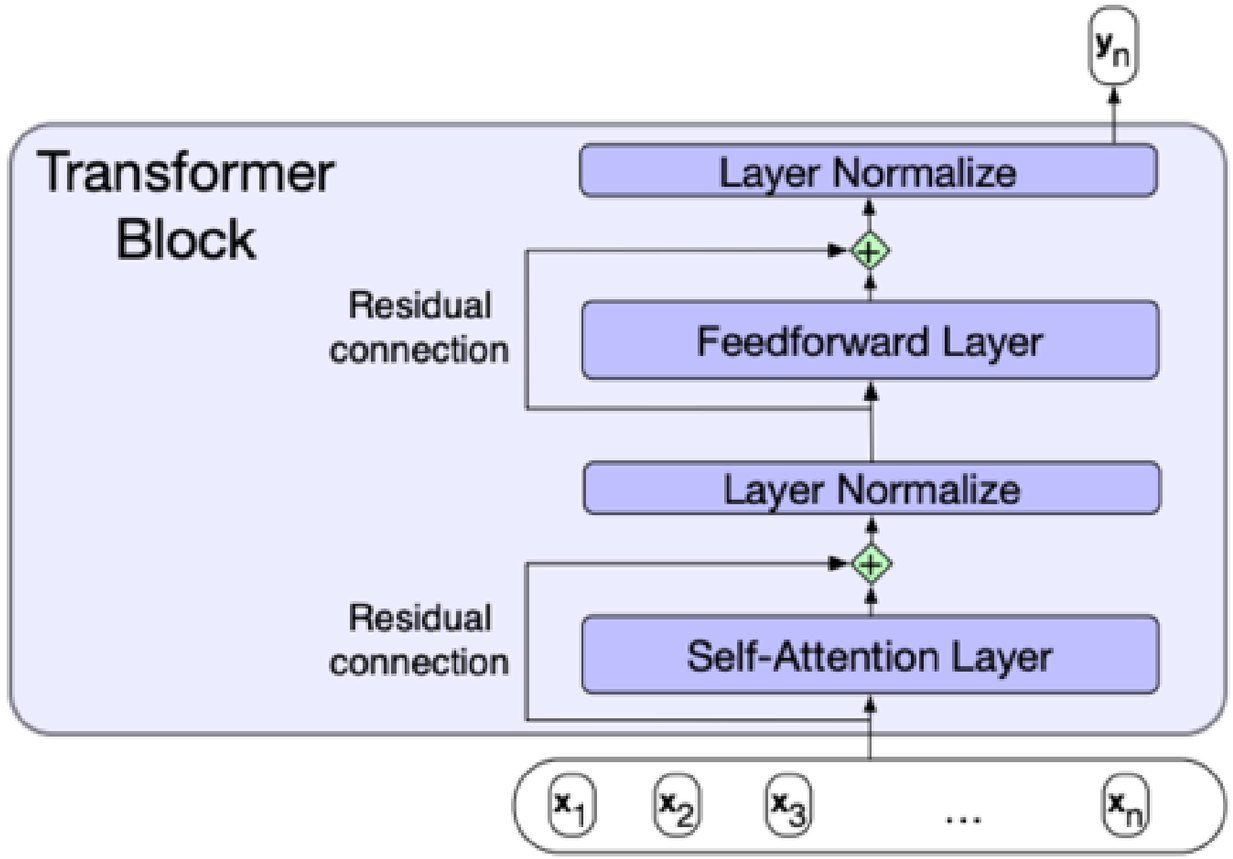
N	q1•k1	−∞	−∞	−∞	−∞
	q2•k1	q2•k2	−∞	−∞	−∞
	q3•k1	q3•k2	q3•k3	−∞	−∞
	q4•k1	q4•k2	q4•k3	q4•k4	−∞
	q5•k1	q5•k2	q5•k3	q5•k4	q5•k5
		N			



ihealth | Millennium Institute for Intelligent Healthcare Engineering

Transformer Blocks

A transformer block showing all the layers

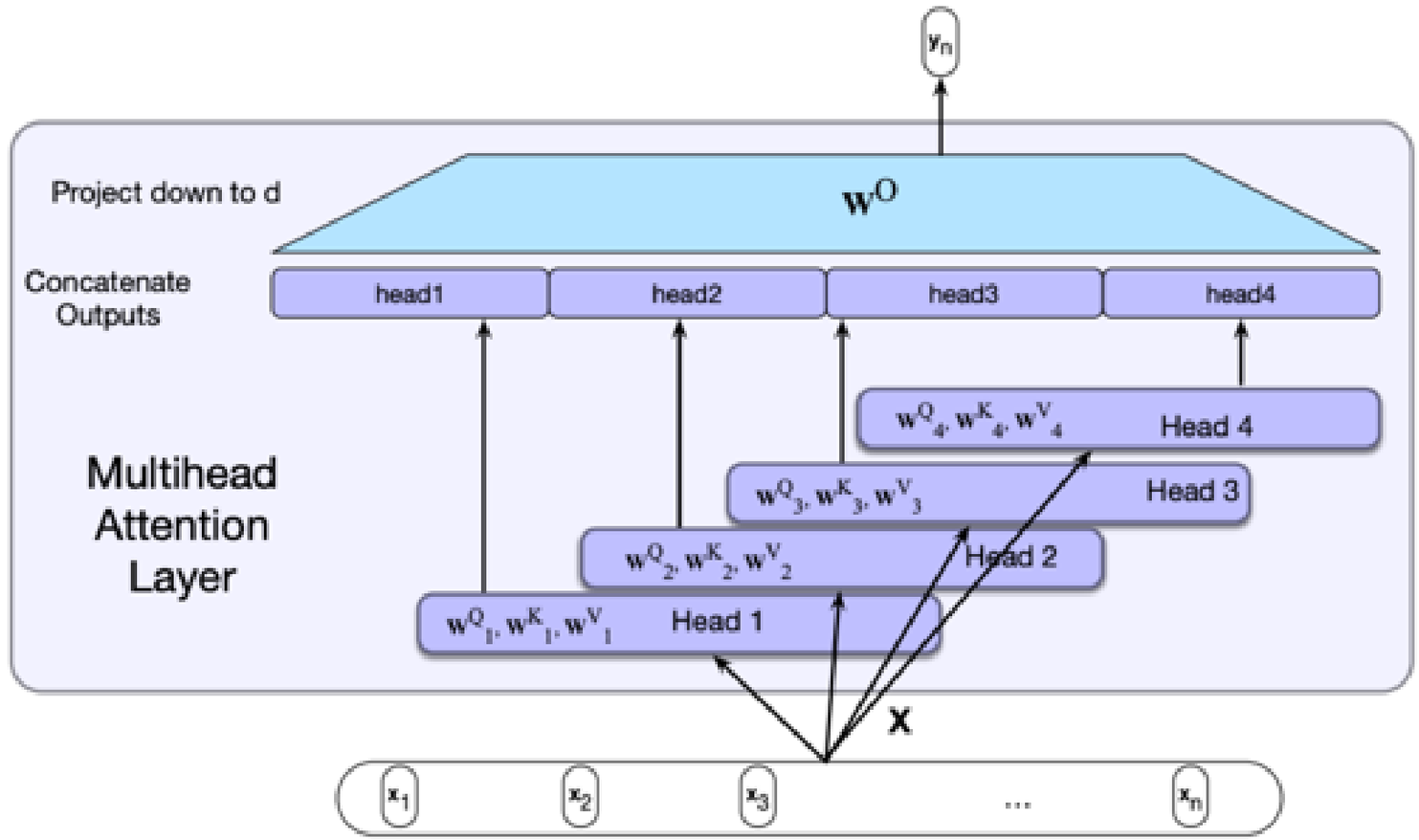


ihealth | Millennium Institute for Intelligent Healthcare Engineering

Multihead Attention

Multihead self-attention: Each of the multihead self-attention layers is provided with its own set of key, query and value weight matrices.

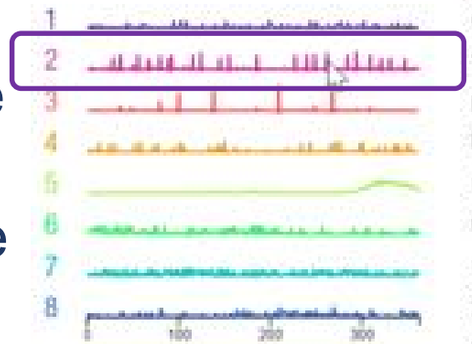
The outputs from each of the layers are concatenated



ihealth | Millennium Institute for Intelligent Healthcare Engineering

Multihead Attention

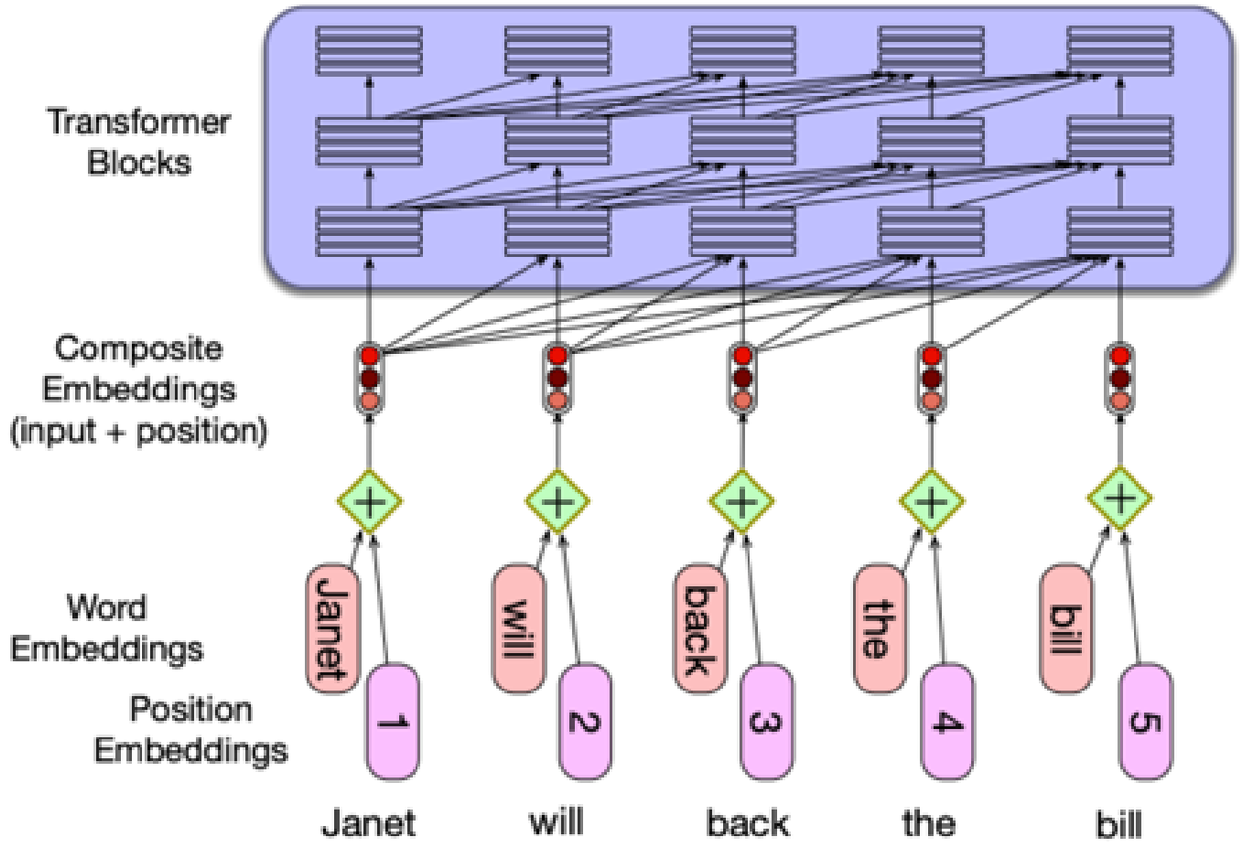
With the python library ecco we can oberve the effect of different heads on processing the same input text



[CLS] now i ask you : what can be expected of man since he is a being endowed with strange qualities ? shower upon him every earthly blessing , drown him in a sea of happiness , so that nothing but bubbles of bliss can be seen on the surface ; give him economic prosperity , such that he should have nothing else to do but sleep , eat cakes and busy himself with the continuation of his species , and even then out of sheer ing rat itude , sheer spite , man would play you some nasty trick . he would even risk his cakes and would deliberately desire the most fatal rubbish , the most une con omic al absurd ity , simply to introduce into all this positive good sense his fatal fantastic element . it is just his fantastic dreams , his vulgar folly that he will desire to retain , simply in order to prove to himself - - as though that were so necessary - - that men still are men and not the keys of a piano , which the laws of nature threaten to control so completely that soon one will be able to desire nothing but by the calendar . and that is not all : even if man really were nothing but a piano - key , even if this were proved to him by natural science and mathematics , even then he would not become reasonable , but would purposely do something per verse out of simple ing rat itude , simply to gain his point . and if he does not find means he will con tri ve destruction and chaos , will con tri ve sufferings of all sorts , only to gain his point ! he will launch a curse upon the world , and as only man can curse (it is his privilege , the primary distinction between him and other animals) , may be by his curse alone he will attain his object - - that is , convince himself that he is a man and not a piano - key ! [SEP]

Modeling positions (sequence)

A simple way to model position: simply adding an embedding representation of the absolute position to the input word embedding.

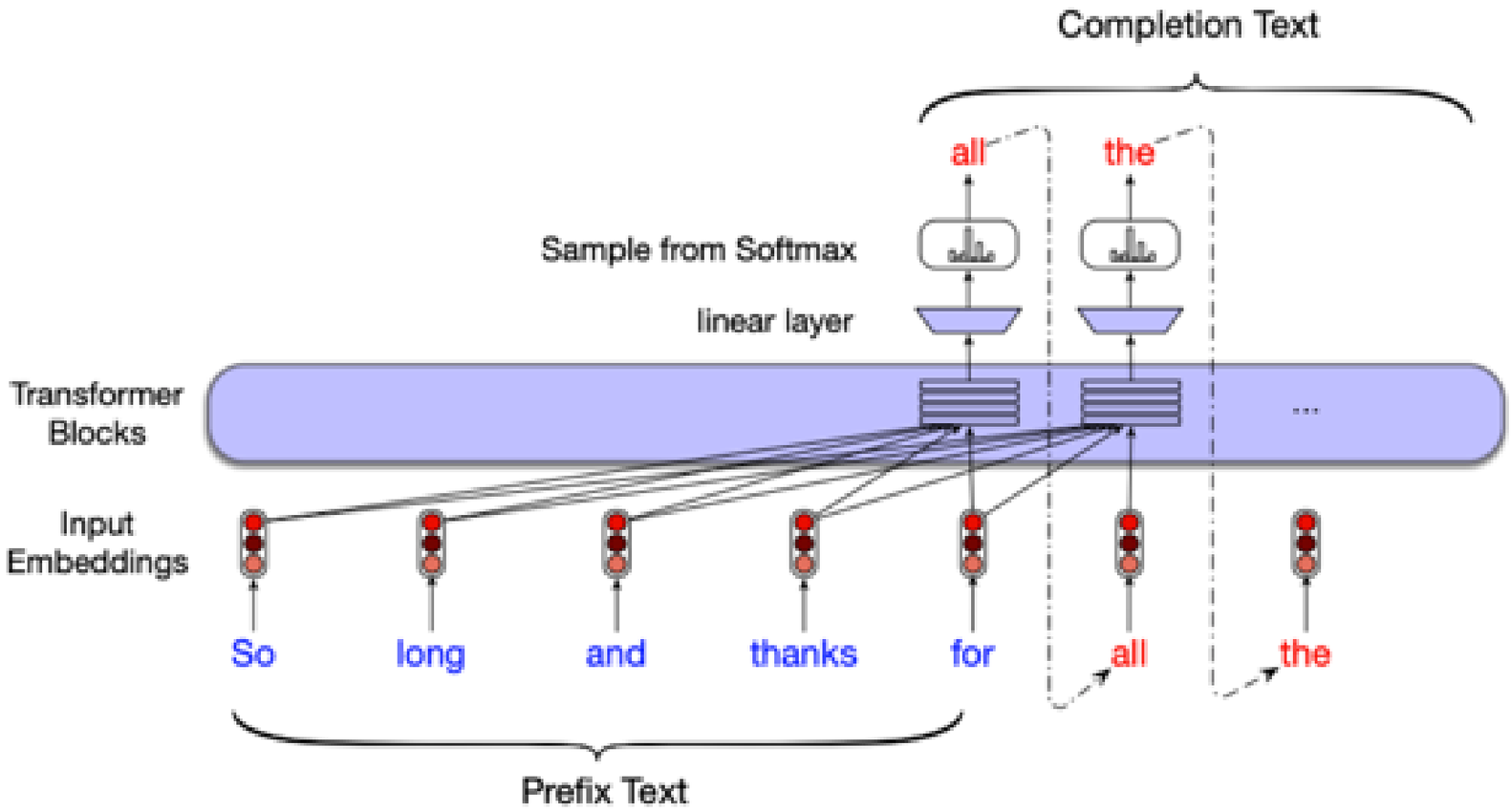


ihealth | Millennium Institute for Intelligent Healthcare Engineering

NLG with Transformers

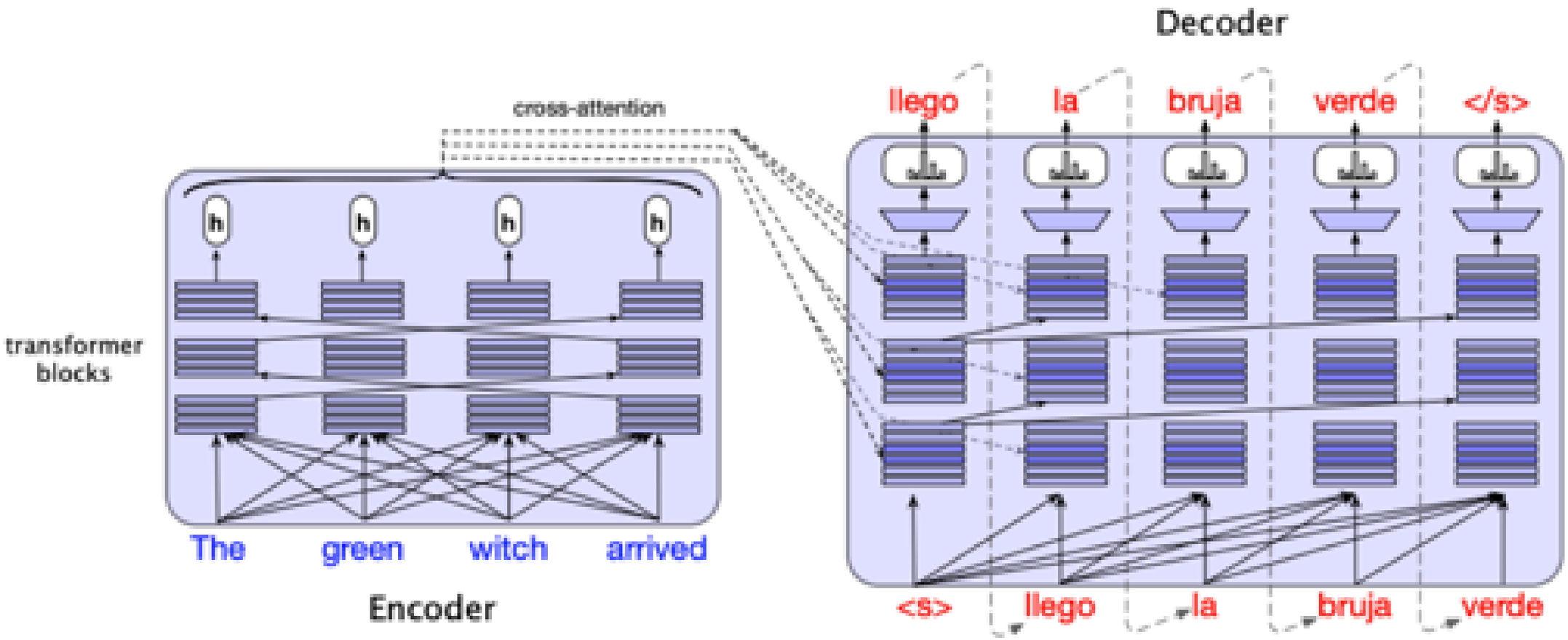
ihealth | Millennium Institute for Intelligent Healthcare Engineering

Text completion



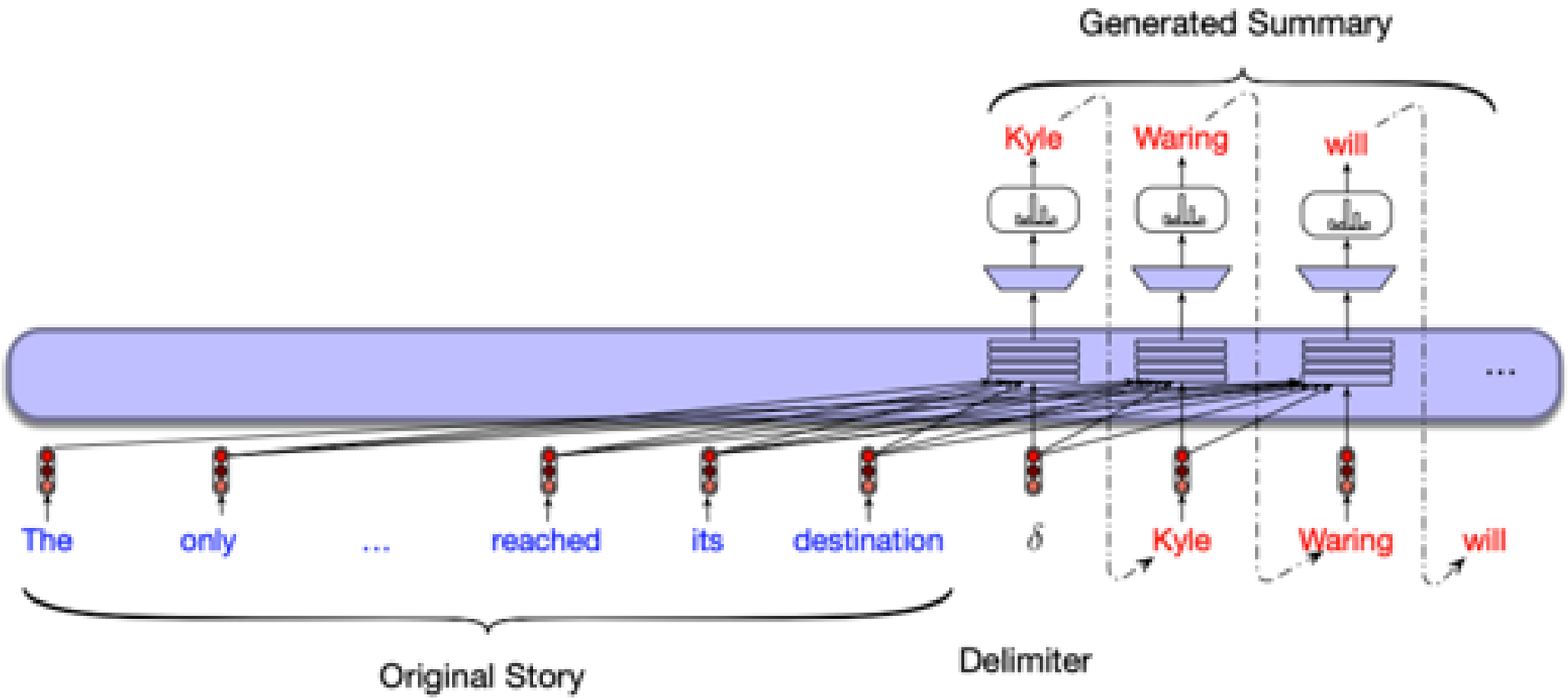
ihealth | Millennium Institute for Intelligent Healthcare Engineering

Machine Translation



ihealth | Millennium Institute for Intelligent Healthcare Engineering

Summarization



ihealth | Millennium Institute for Intelligent Healthcare Engineering

Evaluation metrics (BLEU & ROUGE)

ihealth | Millennium Institute for Intelligent Healthcare Engineering

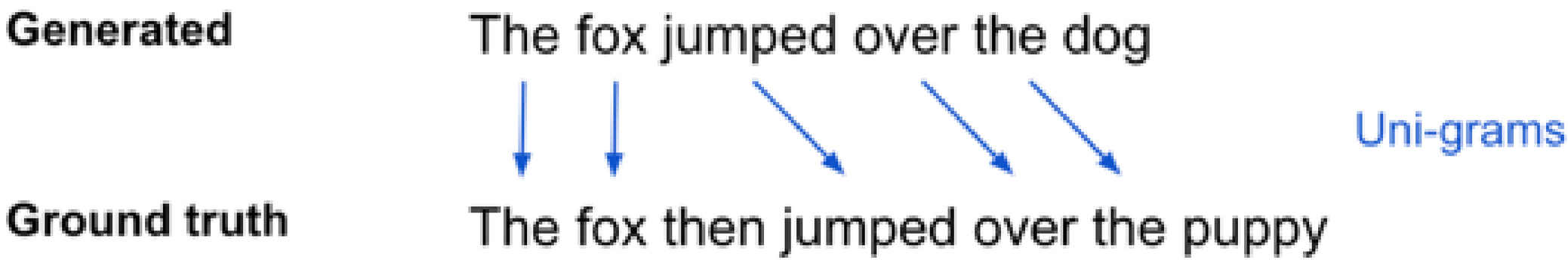
BLEU (Papineni et al., 2002)

- Counts n-grams matches in the ground truth
- Precision-based
- Brevity penalty
- BLEU-N uses up to N-grams (1-4)

ihealth | Millennium Institute for Intelligent Healthcare Engineering

BLEU (Papineni et al., 2002)

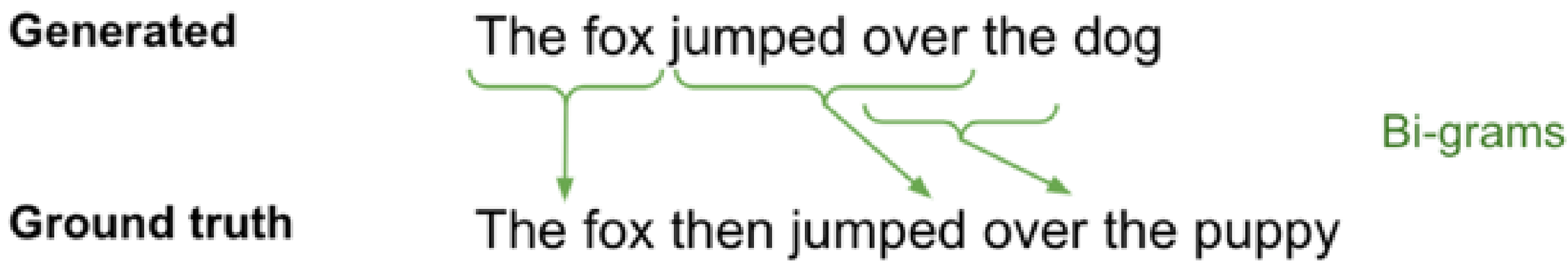
- Counts n-grams matches in the ground truth
- Precision-based
- Brevity penalty
- BLEU-N uses up to N-grams (1-4)



ihealth | Millennium Institute for Intelligent Healthcare Engineering

BLEU (Papineni et al., 2002)

- Counts n-grams matches in the ground truth
- Precision-based
- Brevity penalty
- BLEU-N uses up to N-grams (1-4)



ihealth | Millennium Institute for Intelligent Healthcare Engineering

ROUGE (Lin, 2004)

- Finds the largest common subsequence
- Sentence A is a subsequence of sentence B if:
  - All words of A appear in the same order in B
  - B may have other words in between
- Harmonic average biased toward recall

Generated The fox jumped over the dog

Ground truth The fox then jumped over the puppy

Largest common subsequence

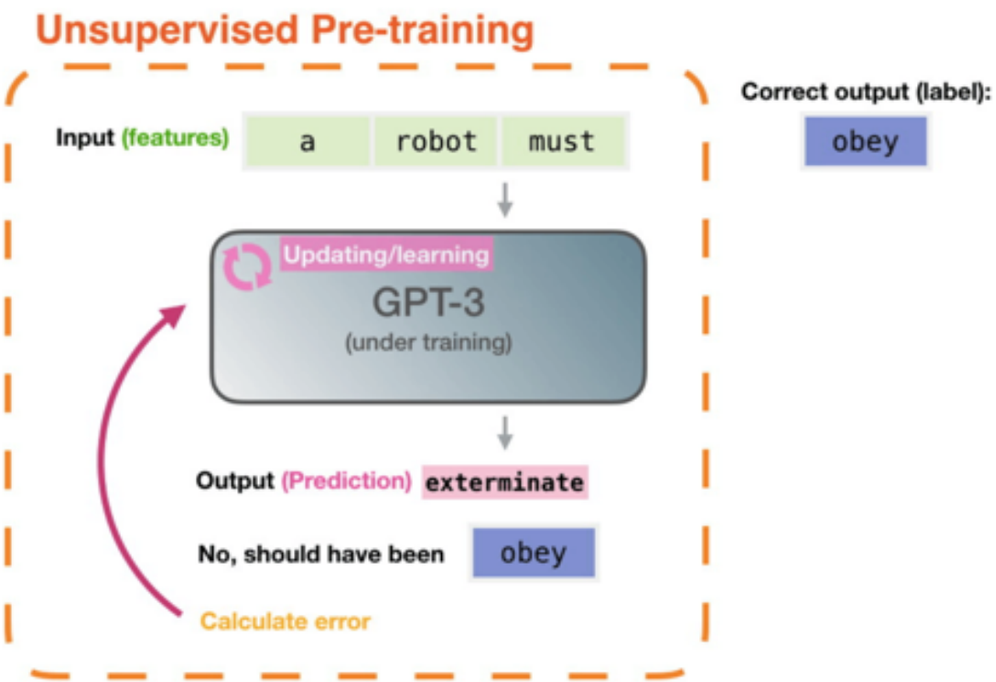
ihealth | Millennium Institute for Intelligent Healthcare Engineering

State of the art in NLG

ihealth | Millennium Institute for Intelligent Healthcare Engineering

GPT3

Generative Pre-Training v3



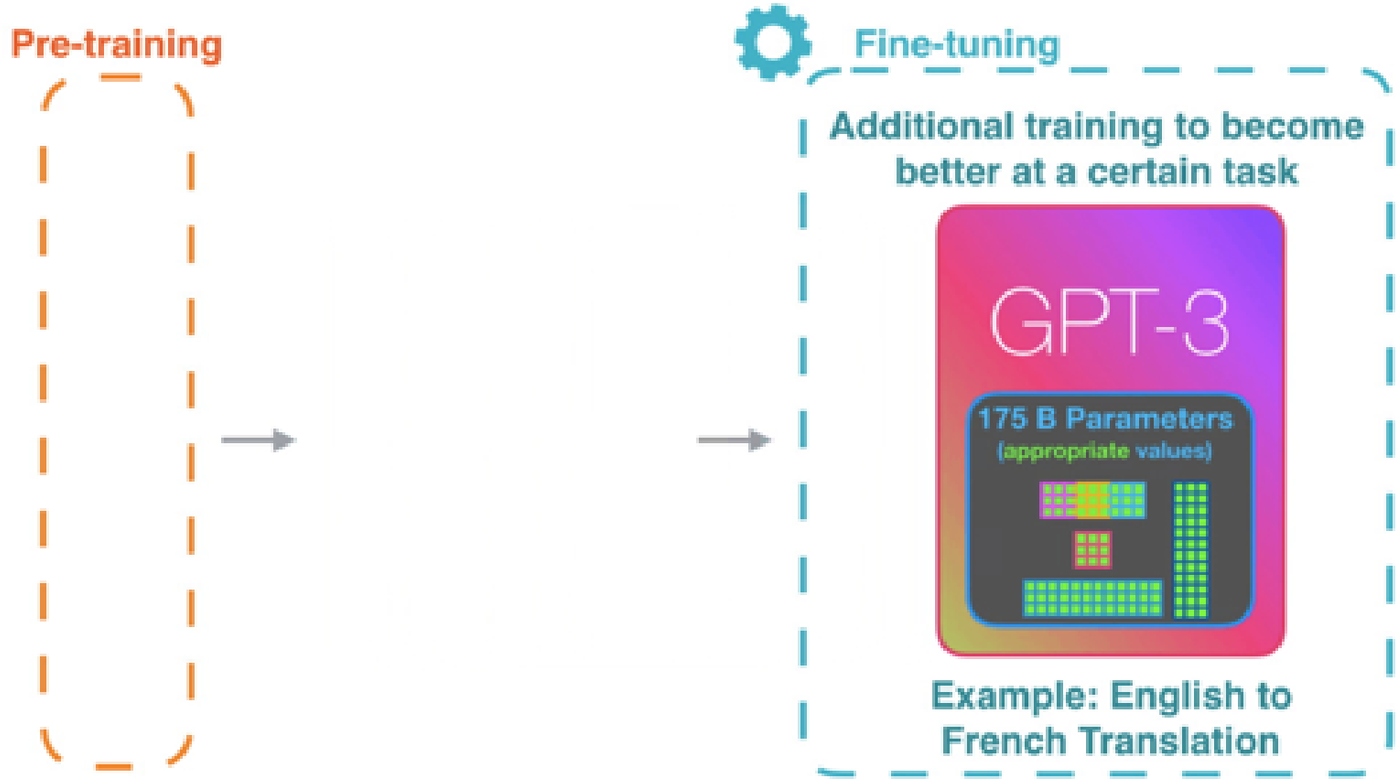
GPT3 is MASSIVE. It encodes what it learns from training in 175 billion parameters.

It was estimated to train in 355 GPU years and cost $4.6m

ihealth | Millennium Institute for Intelligent Healthcare Engineering

GPT3

We can use it to other applications and datasets (e.g. Medical) by finetuning the pre-trained model



ihealth | Millennium Institute for Intelligent Healthcare Engineering

Current models

HUGE # of parameters !

Figure 1: **Parameter counts of several recently released pretrained language models.**

Sanh et al. (2020) DistilBERT, a distilled version of BERT: smaller,faster, cheaper and lighter https://arxiv.org/pdf/1910.01108.pdf

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Current models

Megatron-Turing NLG

We ended with a set of 15 datasets consisting of a total of 339 billion tokens. During training, we opted to blend the datasets into heterogeneous batches according to variable sampling weights given in Figure 2, with an emphasis on higher-quality datasets. We trained the model on 270 billion tokens.

Dataset	Dataset source	Tokens (billions)	Weight (%)	Epochs
Books3	Pile dataset	25.7	14.3	1.5
OpenWebText2	Pile dataset	14.8	19.3	3.6
Stack Exchange	Pile dataset	11.6	5.7	1.4
PubMed Abstracts	Pile dataset	4.4	2.9	1.8
Wikipedia	Pile dataset	4.2	4.8	3.2
Gutenberg (PG-19)	Pile dataset	2.7	0.9	0.9
BookCorpus2	Pile dataset	1.5	1.0	1.8
NIH ExPorter	Pile dataset	0.3	0.2	1.8
Pile-CC	Pile dataset	49.8	9.4	0.5
ArXiv	Pile dataset	20.8	1.4	0.2
GitHub	Pile dataset	24.3	1.6	0.2
CC-2020-50	Common Crawl (CC) snapshot	68.7	13.0	0.5
CC-2021-04	Common Crawl (CC) snapshot	82.6	15.7	0.5
RealNews	RealNews	21.9	9.0	1.1
CC-Stories	Common Crawl (CC) stories	5.3	0.9	0.5

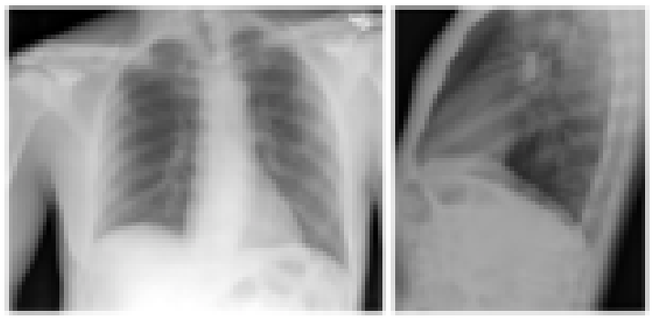
Figure 2. Datasets used to train the MT-NLG model.

Megatron NLG https://www.microsoft.com/en-us/research/blog/using-deepspeed-and-megatron-to-train-megatron-turing-nlg-530b-the-worlds-largest-and-most-powerful-generative-language-model/

ihealth | Millennium Institute for Intelligent Healthcare Engineering

NLG in chest x-rays

In the case of Chest X-rays, we aim to learn a mapping from one or more images to the “findings” text section

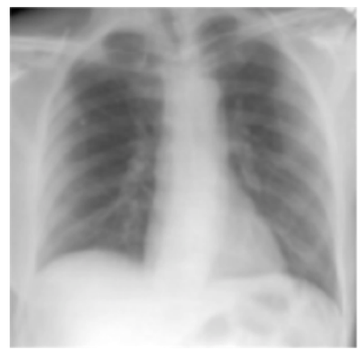


Manual tags: Calcified Granuloma/lung/upper lobe/right
Automatic tags: Calcified granuloma

Comparison: Chest radiographs XXXX.
Indication: XXXX-year-old male, chest pain.
Findings: The cardiomediastinal silhouette is within normal limits for size and contour. The lungs are normally inflated without evidence of focal airspace disease, pleural effusion, or pneumothorax. Stable calcified granuloma within the right upper lung. No acute bone abnormality.
Impression: No acute cardiopulmonary process.

ihealth | Millennium Institute for Intelligent Healthcare Engineering

In our case, we want LMs to predict:

$$P(\textit{granuloma within right lung} \mid \text{  })$$

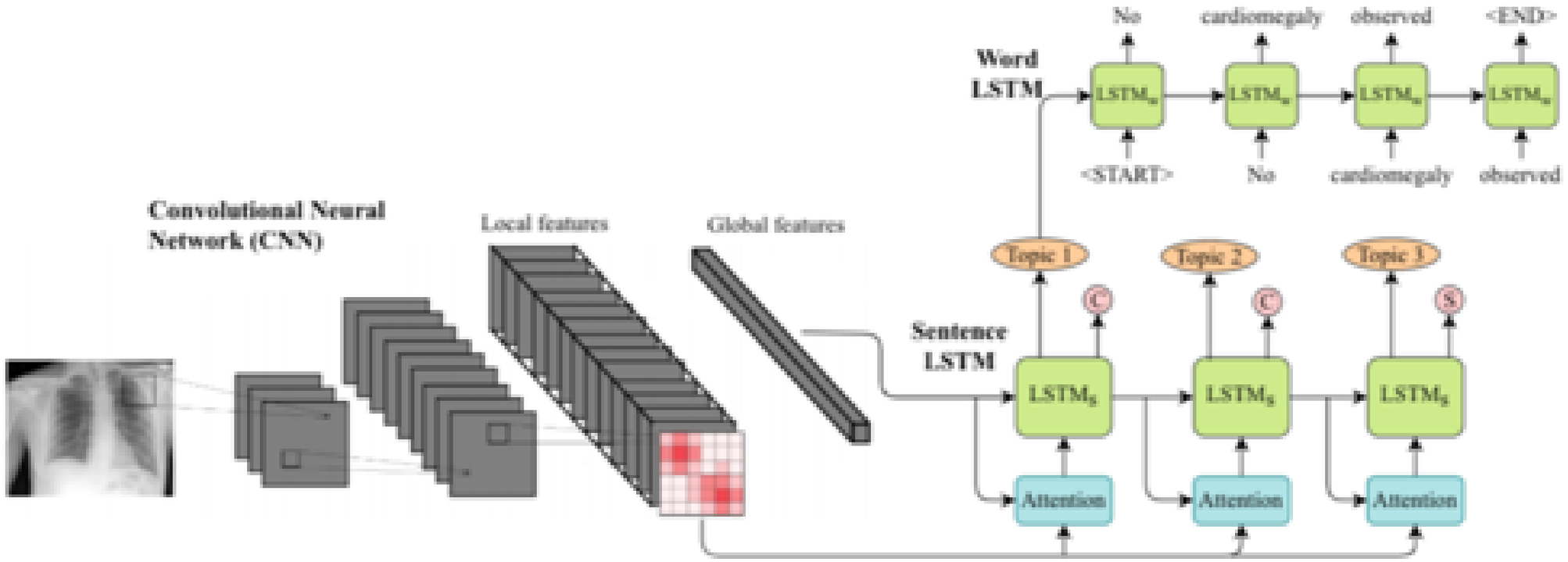
???

ihealth | Millennium Institute for Intelligent Healthcare Engineering

How do we generate text by making the computer “observe” the image ?

ihealth | Millennium Institute for Intelligent Healthcare Engineering

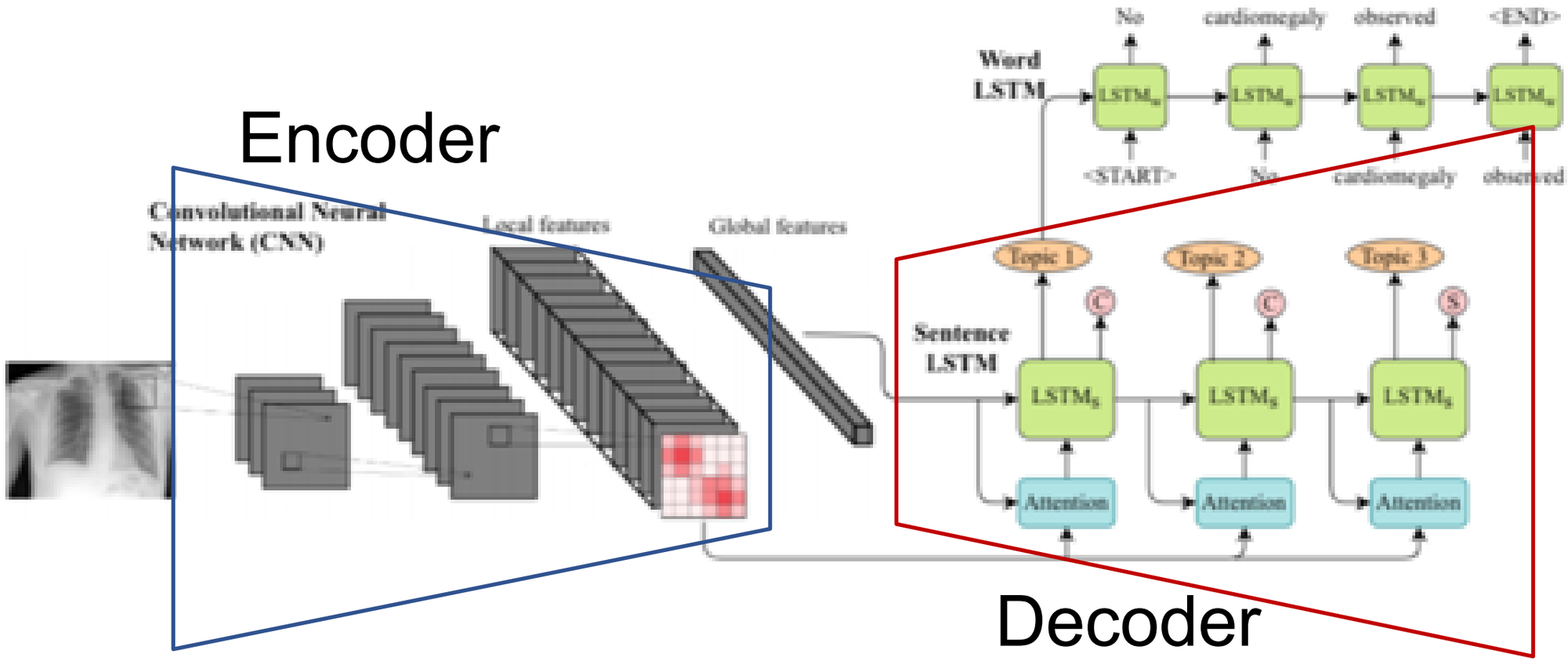
We need to combine architectures and representations!



Messina, P, Pino, P et al. (2020) “A Survey on Deep Learning and Explainability for Automatic Report Generation from Medical Images”

ihealth | Millennium Institute for Intelligent Healthcare Engineering

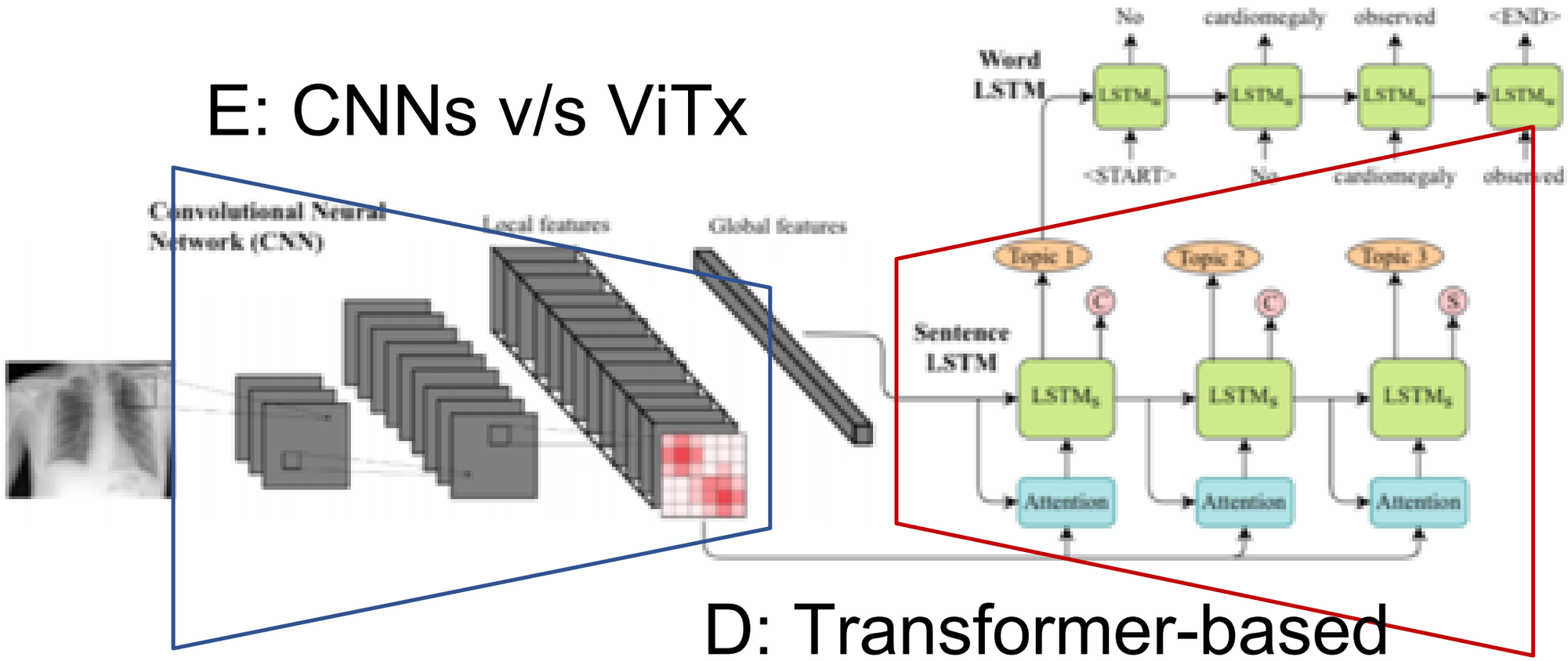
We need to combine architectures and representations!



Messina, P, Pino, P et al. (2020) "A Survey on Deep Learning and Explainability for Automatic Report Generation from Medical Images"

ihealth | Millennium Institute for Intelligent Healthcare Engineering

We need to combine architectures and representations!



Messina, P, Pino, P et al. (2020) "A Survey on Deep Learning and Explainability for Automatic Report Generation from Medical Images"

ihealth | Millennium Institute for Intelligent Healthcare Engineering

A lot of very challenging research ahead!

- Expand to other pathologies and types of images (MRI, CT-Scan, Ecography, etc.)
- Deal with multimodal input (text, images, videos, tabular)
- Interpretable AI: explain predictions of results
- Improve generalization (OOD samples): transfer learning, metalearning, etc.
- Evaluation: NLG metrics vs actual clinical diagnostic

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Now is your turn!

- Hands on session with Pablo Messina
- 3 Google Colab Notebooks

https://github.com/PabloMessina/Code-for-SIPAIM-2022-NLG-Tutorial

ihealth | Millennium Institute for Intelligent Healthcare Engineering

Acknowledgment

This work was funded by ANID - Millennium Science Initiative Program - ICN2021_004