



Recap

Liad Magen

NLP - Usages

- Rule-based information extraction (RegExp)
- Classification
 - Sentence boundary
 - Ambiguity (PP attachment problem)
 - Sentiment/Opinion/Toxic texts
 - Named Entity Recognition (NER)
 - Semantic Role Labeling (SRL) / Part of speech (POS)

Methods

One-hot-vector + TF/IDF

Naïve Bayes / SVM

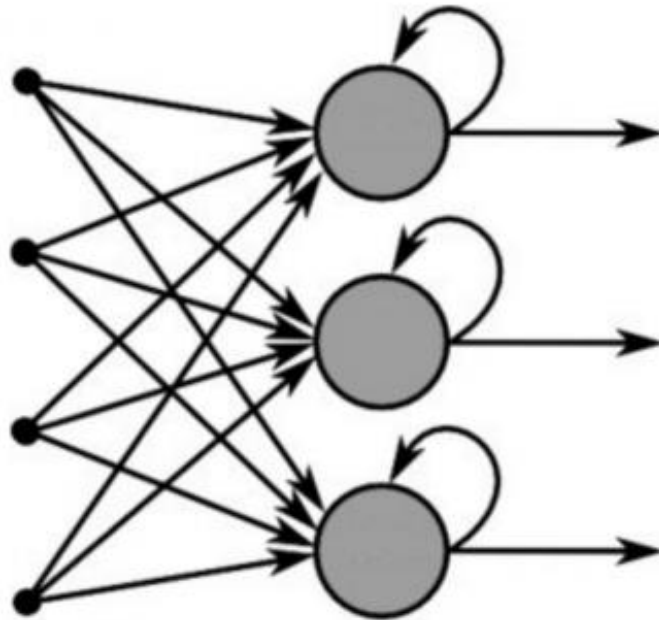
Word Embedding (GloVe / FastText)

BI-LSTM

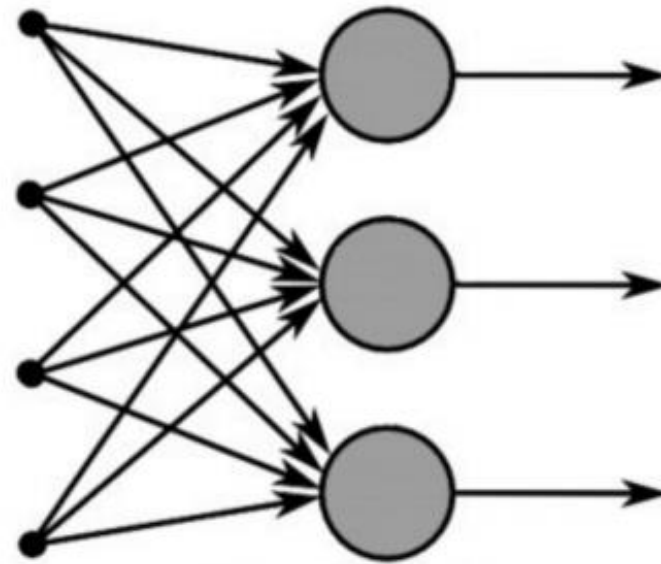
Contextual Word Embedding (ELMo)

Transformers / BERTology

RNN

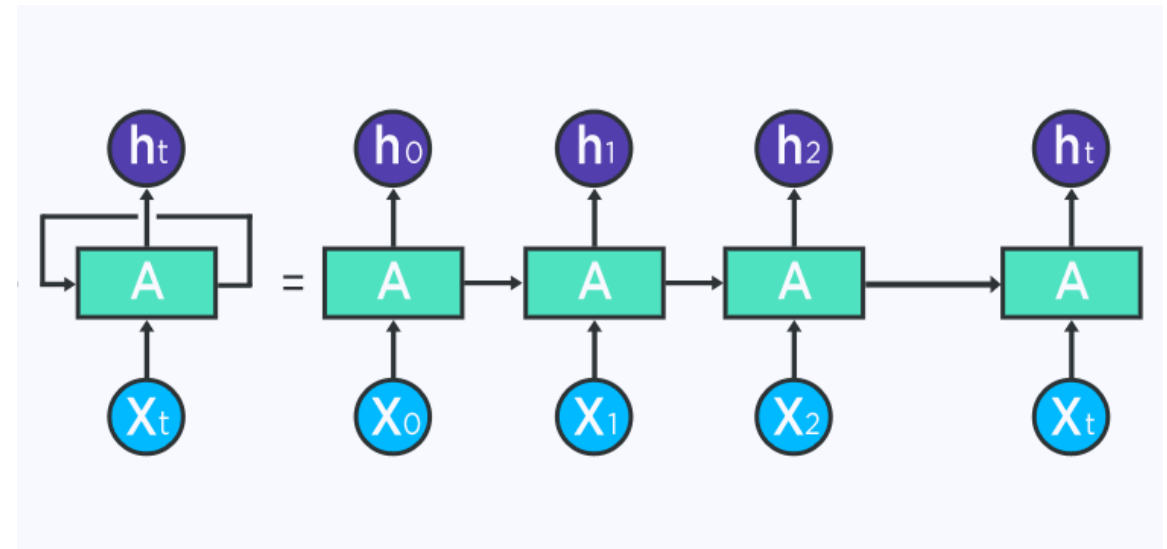
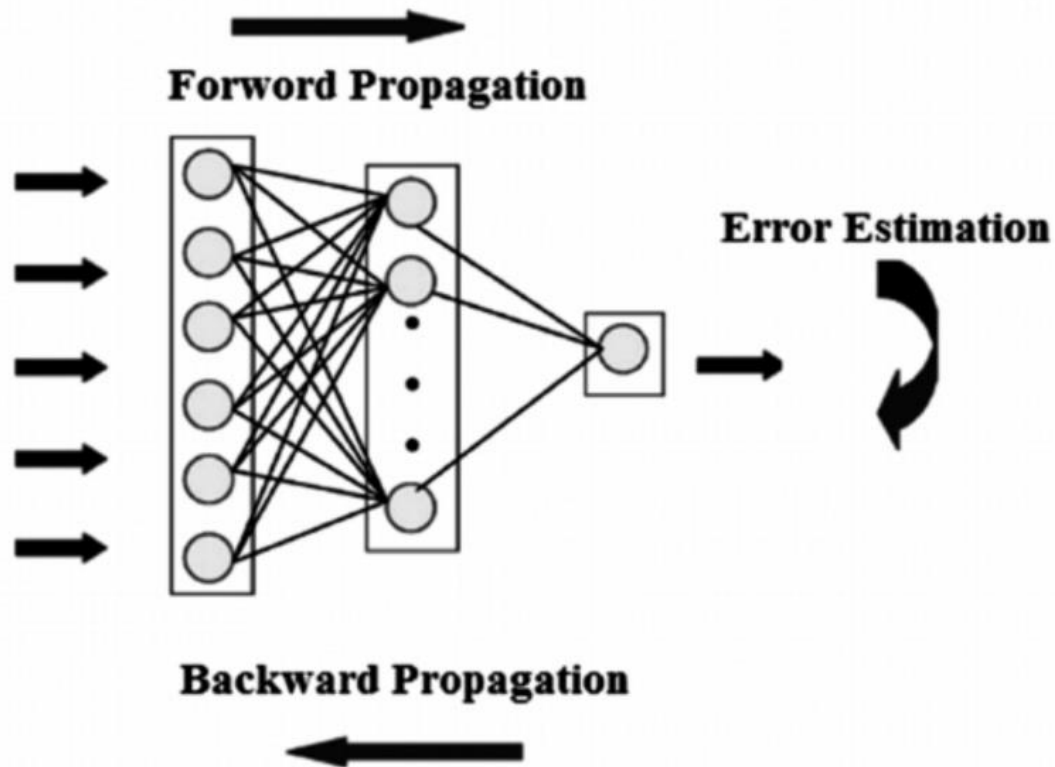


Recurrent Neural Network

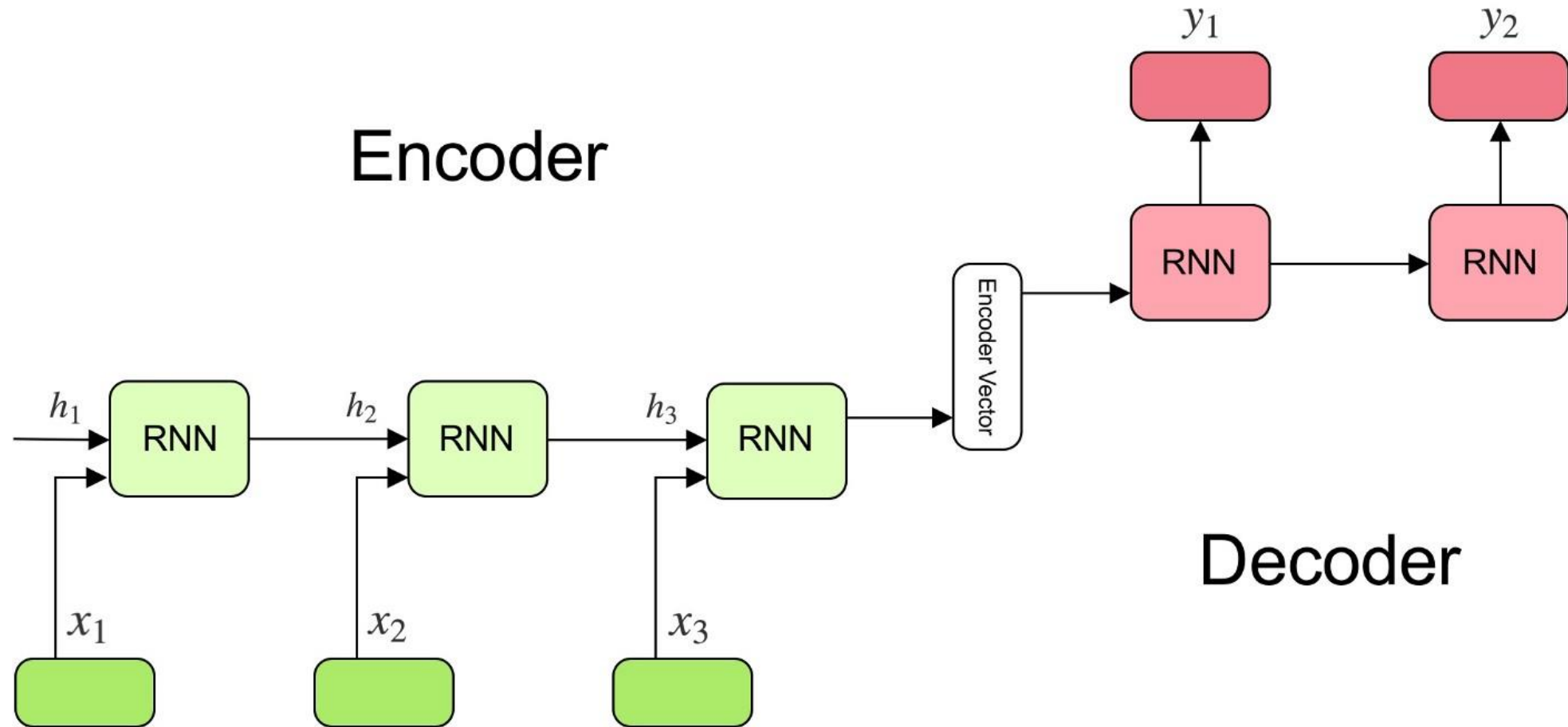


Feed-Forward Neural Network

RNN – Back Propagation

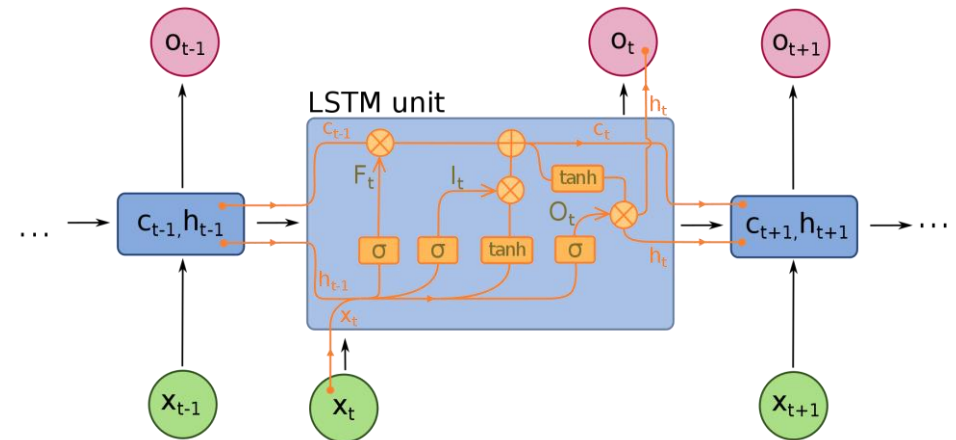


RNN – Encoder/Decoder

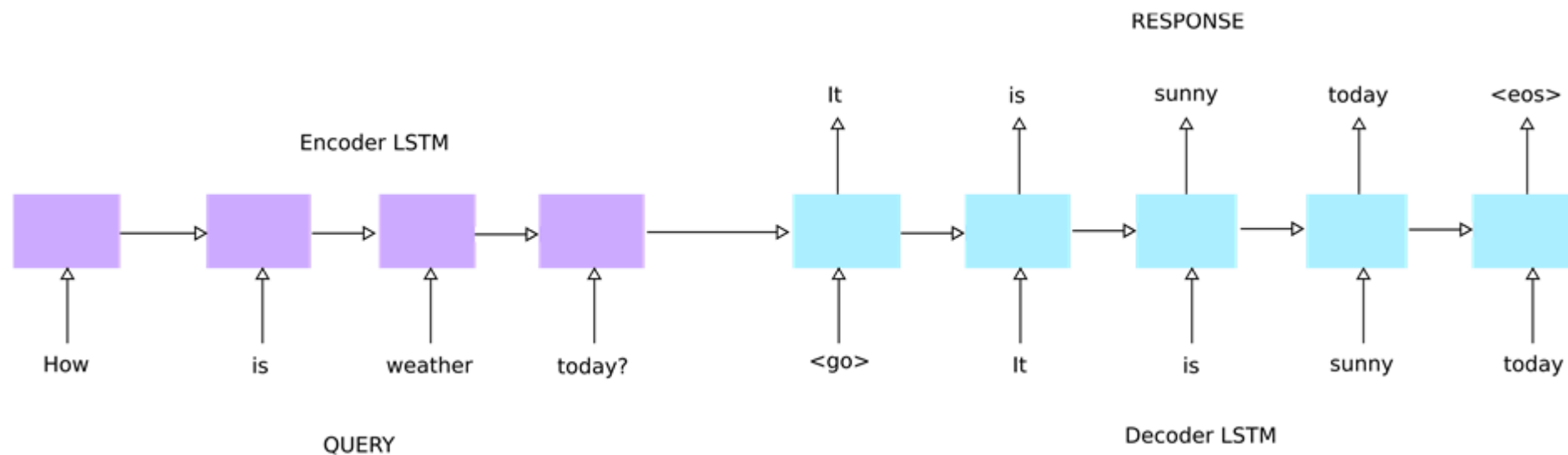


LSTM

- A better version of RNN
 - Easier to train
 - Overcomes RNN issue of 'forgetting' and 'not learning' (vanishing gradient)
- BI-LSTM – Bi-directional LSTM
 - Can 'peek' into the future

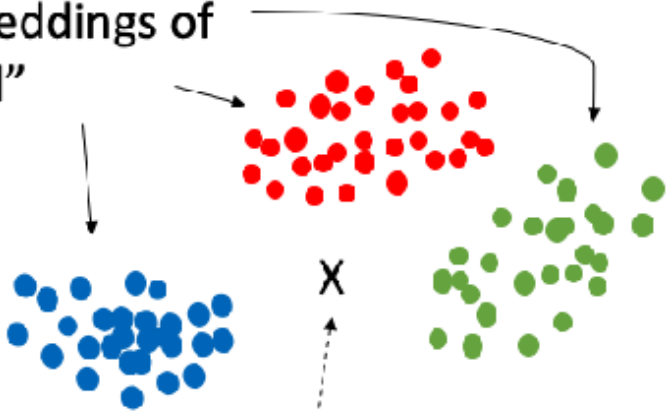


LSTM – Encoder/Decoder (Seq2Seq)



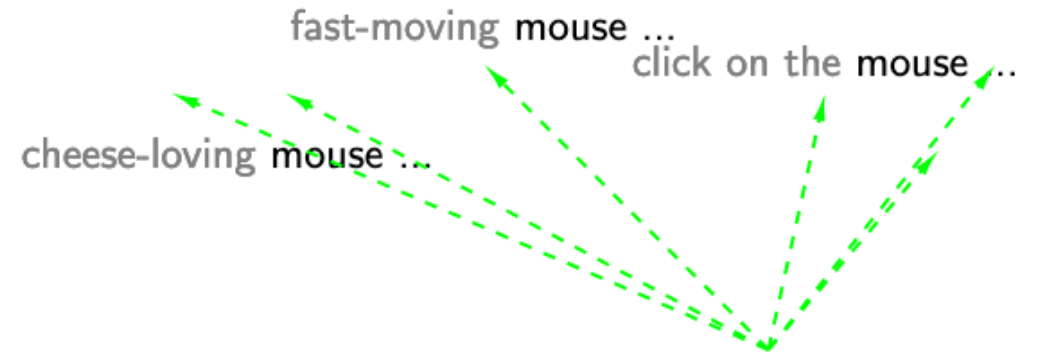
Contextual Word Embedding

Contextualized
embeddings of
"cold"



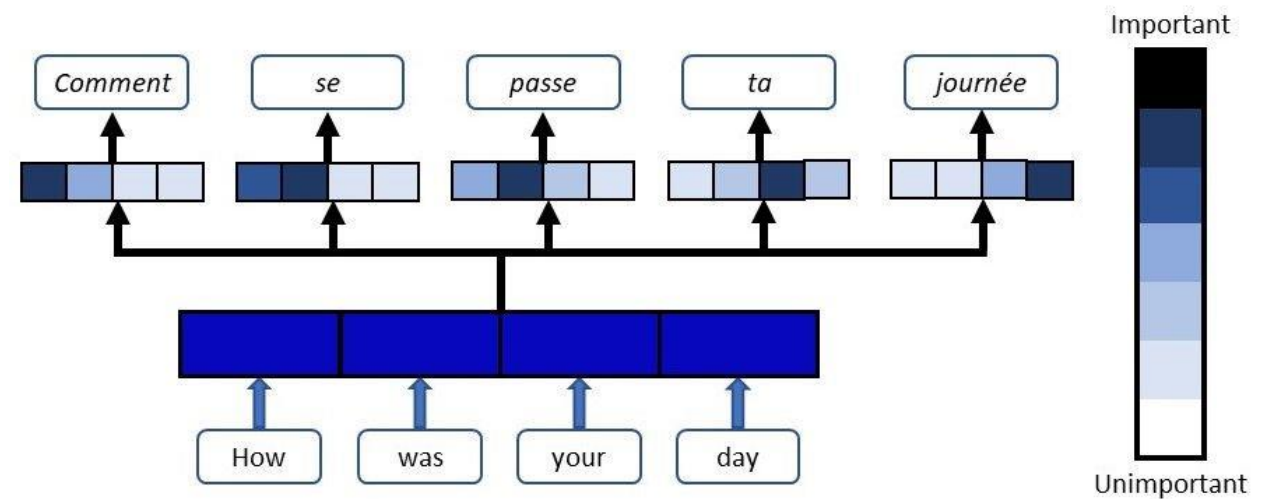
Word
embedding
of "cold"

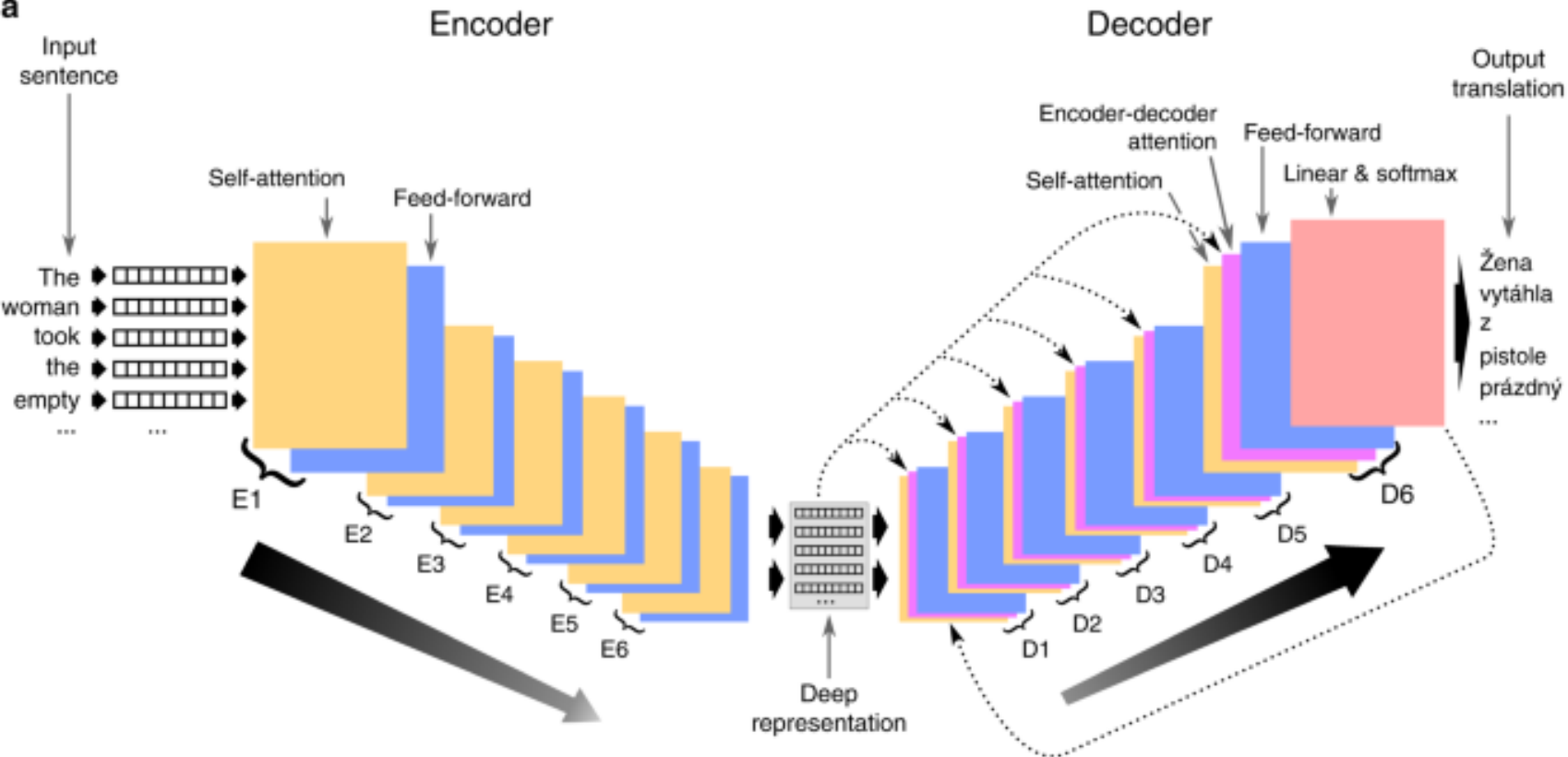
cold (temperature)
cold (symptom)
cold (unfriendly)



Attention Mechanism

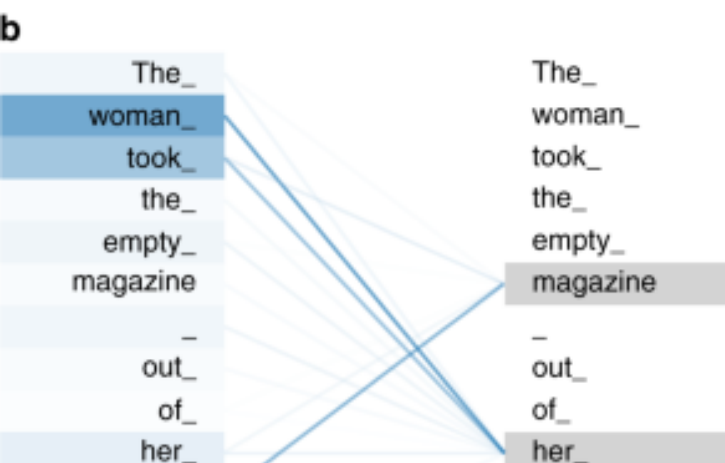
- An architecture to learn the importance of words (X) and their inter-relations to a given target Y
- Learned Matrix

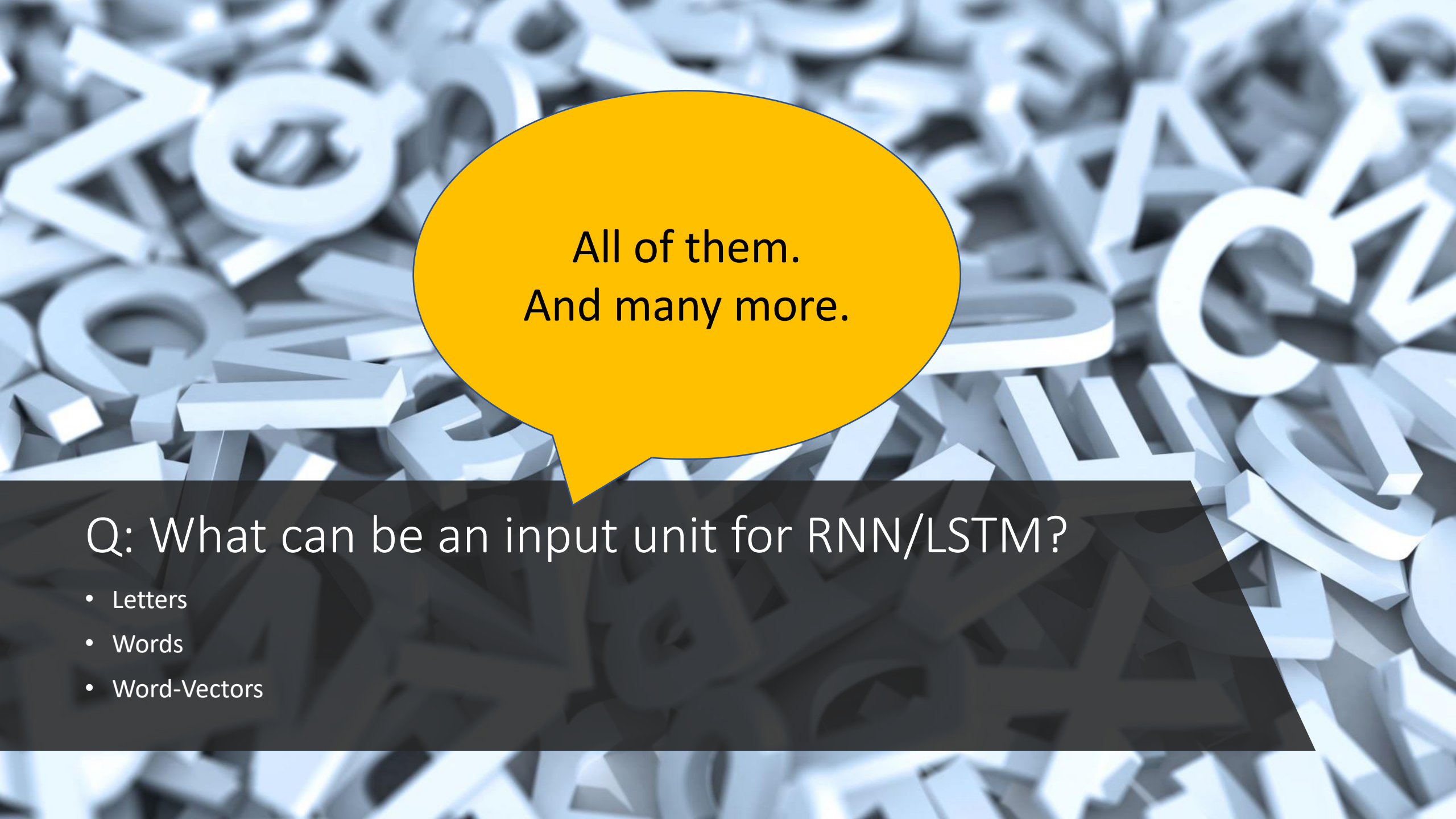




Transformers

- Encoder-Decoder
- Multiple Attention Matrices



The background of the slide is a dense, out-of-focus field of 3D, light-blue letters and numbers. In the center, there is a bright yellow speech bubble with a black outline. Inside the bubble, the text "All of them. And many more." is written in a black, sans-serif font.

All of them.
And many more.

Q: What can be an input unit for RNN/LSTM?

- Letters
- Words
- Word-Vectors

Q: What is
the input unit
for BERT?

