Lecture

# Knowledge-based Systems

# Part 4 – Pre-trained Language Models

**Dr. Mohsen Mesgar**

**Universität Duisburg-Essen**

# Exam

- In total 21 students have participated in the survey.
- The exam date is **01.08.2022 16:00 -18:00.**
- Die globale **Anmeldephase** läuft vom **02.05.2022** bis **13.05.2022**
- **Where?** I'll update you

Pooling exam date

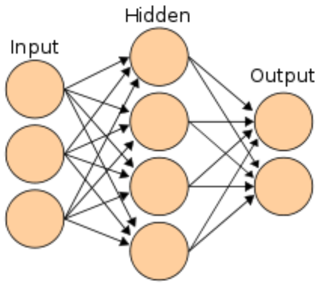
1 Check all dates in which you can take the exam. The exam is schriftlich and takes 2 hours.

| Response | Average | Total |
|---|---|---|
| 01.08.2022 16:00-18:00 | 62% | 13 |
| 02.08.2022 16:00-18:00 | 57% | 12 |
| 03.08.2022 10:00-12:00 | 48% | 10 |
| | | |
| Total responses to question | 100% | 21/21 |

# Recall …

UNIVERSITÄT DUISBURG ESSEN
*Offen im Denken*

- **What is (artificial) intelligence?** The ability to acquire and apply knowledge and skills to achieve complex goals.
- **Symbolic**: Knowledge is encoded by symbols that refer to the knowledge.
- **connectionist**: Knowledge is **embedded** in parameters of a model.



# Any other open questions?

# In this lecture, you learn about …

- **Pretrained language models (LMs)**
  - Unidirectional
  - Bidirectional
- **LMs as knowledge base**
  - LMs and factual knowledge
  - LMs and linguistic knowledge
  - LMs and word sense knowledge

# Unidirectional Language Models

UNIVERSITÄT DUISBURG ESSEN

*Offen im Denken*

- Given an input sequence of tokens **w = [$w_1$,$w_2$,...,$w_N$],** unidirectional language models assign a probability ***p*(w)** to the sequence.
- This probability is calculated as follows

$$p(\mathbf{w}) = \prod_t p(w_t \,|\, w_{t-1}, \ldots, w_1).$$

# Example

$$P_{(w_1, w_2, ..., w_n)} = p(w_1)p(w_2|w_1)p(w_3|w_1, w_2)...p(w_n|w_1, w_2, .., w_{n-1})$$
$$= \prod_{i=1}^{n} p(w_i|w_1, ..., w_{i-1})$$

S = Where are we going

Previous words (Context)

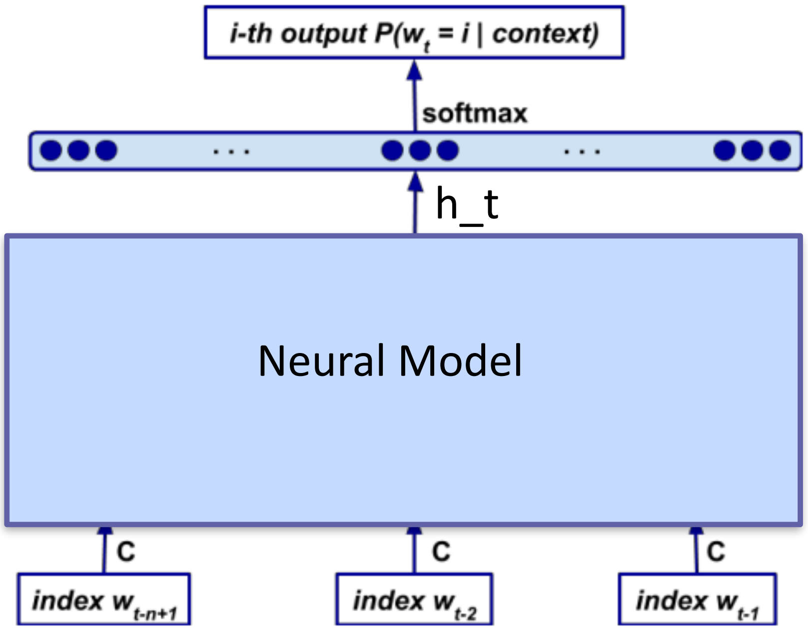
Word being predicted

P(S) = P(Where) x P(are | Where) x P(we | Where are) x P(going | Where are we)

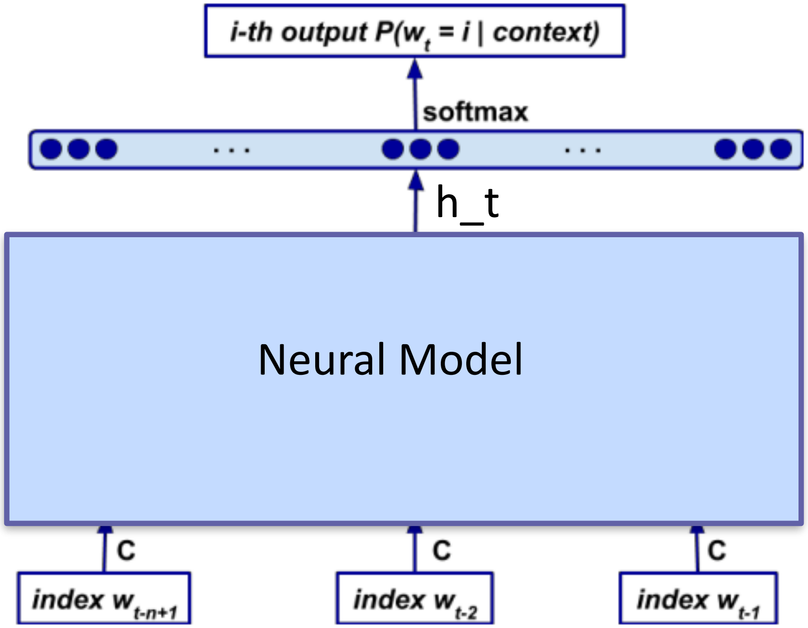
https://thegradient.pub/understanding-evaluation-metrics-for-language-models/

# How to get the probability?

- There are different ways to define the probability function
  - p(w_t|w_(t_1),…w_1)
- State-of-the-art LMs use deep neural models and softmax to estimate the probability



# More formally



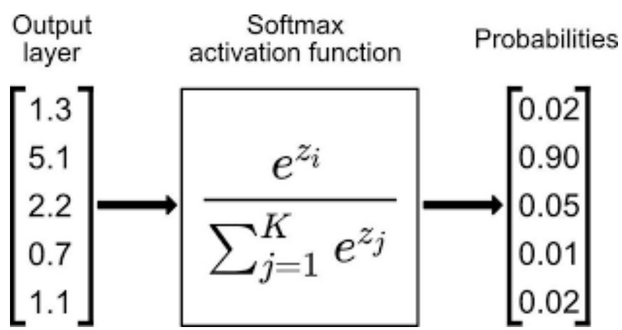

$$p(w_t \,|\, w_{t-1}, \ldots, w_1) = \mathrm{softmax}(\mathbf{W}\mathbf{h}_t + \mathbf{b})$$

parameter of the output layer

output vector of a neural network at position *t*

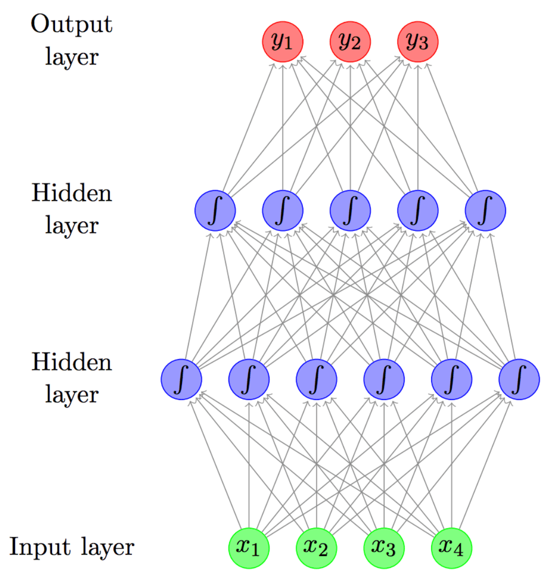
# Softmax

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$



# Knowledge is embedded

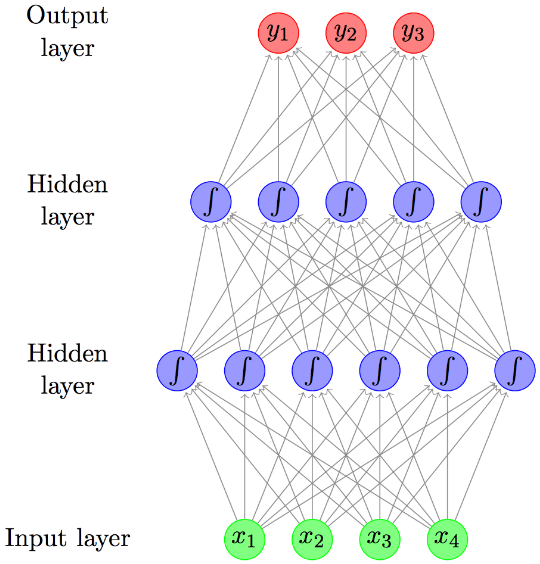
- The knowledge about words and their relations in a language is encoded in the parameters (connections) of the neural language model



# Knowledge is embedded



- The knowledge about words and their relations in a language is encoded in the parameters (connections) of the neural language model




Today, we assume that the model already knows the knowledge. The model is **pretrained.**
*"How to train LMs"* is what we discuss in next lectures.

# Architecture of Neural Language Models

- The difference in the neural language models is in how they compute h_t
- Different architectures have been explored
  - Multi-layer-perceptron
  - Convolutional layers
  - Recurrent neural networks
  - Transformers (self-attention mechanism)

# Examples of unidirectional LM

UNIVERSITÄT DUISBURG ESSEN

*Offen im Denken*

- **Fairseq-fconv** (http://proceedings.mlr.press/v70/dauphin17a.html)
  - Convolutional neural model

- **Transformer-XL** (https://arxiv.org/abs/1901.02860)
  - Transformer-based model

# Bidirectional Language Models

UNIVERSITÄT DUISBURG ESSEN

*Offen im Denken*

- In many downstream applications we mostly care about having access to contextual representations of words,
- word representations are a function of the **entire context** of a unit of text such as a sentence or paragraph, and not only conditioned on previous words.

$$p(w_i) = p(w_i \mid w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_N)$$

# Examples of Bidirectional LM

- **ELMo** (https://allenai.org/allennlp/software/elmo)
  
  - Deep RNN-based LM
  - At each layer, one LSTM processes words from left to right, the other processes words from right to left)
- **BERT**
  
  - Transformer-based LM
  - Uses self-attention mechanism to condition representations of a word on its le and right context
  
- **BART**
  - Transformer-based LM
- **RoBERTa**
  - Transformer-based LM
- **GPT**
  - Transformer-based LM

# Practice I

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

- Use google Colab (https://colab.research.google.com)
  - More information (https://huggingface.co/course/chapter0/1)
- Try out 20 different contexts to see what words BERT suggests for the next word
  - https://rb.gy/3k5bsc

# World Knowledge

- We observed that symbolic KB can give us factual knowledge about world
- **Google RE:** place_of_death, date_of_birth, education_degree, place_of_birth (https://code.google.com/archive/p/ relation- extraction-corpus/)

# LM and factual knowledge

- Define a template to query LMs
  - place_of_death —> [S] died in [O]
- 

```
result = unmasker(" Diego de Arroyo died in [MASK].")
print([r["token_str"] for r in result])
```

```
['madrid', 'manila', 'lima', 'seville', 'barcelona']
```

# Practice II

- Use your notebook in Google Colab (https://colab.research.google.com)
- Download the **Google RE dataset** (https://code.google.com/archive/p/relation-extraction-corpus/)
  - Focus on "**place of birth**", "**date of birth**" and "**place of death**" relations
  - **How many facts do exist for each relation?**
- Define **a template for each relation** to query a LM
- Select a LM, e.g. BERT, RoBERTA, ELMo, …
- **For how many facts does the selected LM return the correct value?**
  - **compute P@1**
  - P@k: Is the correct value among the k top outputs that the LM returns?
- Write a report in overleaf **without screen shots**

# LMs and commonsense relationships between words

- **ConceptNet**
  - a multi- lingual knowledge base,
  - built on top of Open Mind Common Sense (OMCS) sentences
  - OMCS represents commonsense relationships be- tween words and/or phrases
  - English part of ConceptNet has single-token objects covering 16 relations
  - For this knowledge source there is no explicit alignment of facts to Wikipedia sentences.

# LMs and commonsense relationships between words



- **ConceptNet**

| | | | | |
|---|---|---|---|---|
| ConceptNet | AtLocation | You are likely to find a overflow in a ____. | drain | sewer [-3.1] , canal [-3.2] , toilet [-3.3] , stream [-3.6] , **drain [-3.6]** |
| | CapableOf | Ravens can ____. | fly | **fly [-1.5]** , fight [-1.8] , kill [-2.2] , die [-3.2] , hunt [-3.4] |
| | CausesDesire | Joke would make you want to ____. | laugh | cry [-1.7] , die [-1.7] , **laugh [-2.0]** , vomit [-2.6] , scream [-2.6] |
| | Causes | Sometimes virus causes ____. | infection | disease [-1.2] , cancer [-2.0] , **infection [-2.6]** , plague [-3.3] , fever [-3.4] |
| | HasA | Birds have ____. | feathers | wings [-1.8] , nests [-3.1] , **feathers [-3.2]** , died [-3.7] , eggs [-3.9] |
| | HasPrerequisite | Typing requires ____. | speed | patience [-3.5] , precision [-3.6] , registration [-3.8] , accuracy [-4.0] , **speed [-4.1]** |
| | HasProperty | Time is ____. | finite | short [-1.7] , passing [-1.8] , precious [-2.9] , irrelevant [-3.2] , gone [-4.0] |
| | MotivatedByGoal | You would celebrate because you are ____. | alive | happy [-2.4] , human [-3.3] , **alive [-3.3]** , young [-3.6] , free [-3.9] |
| | ReceivesAction | Skills can be ____. | taught | acquired [-2.5] , useful [-2.5] , learned [-2.8] , combined [-3.9] , varied [-3.9] |
| | UsedFor | A pond is for ____. | fish | swimming [-1.3] , fishing [-1.4] , bathing [-2.0] , **fish [-2.8]** , recreation [-3.1] |

# Practice III

UNIVERSITÄT DUISBURG ESSEN

*Offen im Denken*

```python
result = unmasker("Birds have [MASK].")
print([r["token_str"] for r in result])
```

```
['wings', 'eyes', 'feathers', 'nectar', 'nests']
```

# Linguistics knowledge

- subject-verb agreement in English

```
result = unmasker("the game that the guard hates [MASK] bad .")
print([r["token_str"] for r in result])
```

```
['is', 'was', 'the', 'goes', 'sounds']
```

# Practice IV

- How to get dataset for subject-verb agreement?
  - Go to wikipedia or any other textual corpus in NLTK
  - Extract 1000 sentences
  - Mask all verbs
    - How to automatically find which word is a verb? Use NLTK or SpaCy
- https://github.com/BeckyMarvin/LM_syneval
- For how many sentences your LM returns a verb that is in agreement with its subject? Report P@1
- Write a paragraph about this experiment in overleaf.

# Linguistics knowledge

- **Anaphora**
  - *“**Tina** went to bed as soon as **she** reached home”,*
    - both Tina and she refer to the same person.
    - Tina is called an “**antecedent**” and she an “**anaphor**”.

# Linguistics knowledge

UNIVERSITÄT DUISBURG ESSEN

*Offen im Denken*

- **Reflexive anaphora**
  - are those that use reflexive pronouns, i.e., pronouns that end in –self or –selves.
  - When a sentence’s subject and object refers to the same individual, we use reflexive anaphora

- *”**Peter** shot **himself** in the foot.”*
- *”**Peter** bounced the ball to **himself**.”*
- *”**Amy and Lizzie** cried **themselves** to sleep.”*

# Linguistics knowledge

- Reflexive Anaphora

```
result = unmasker("Amy and Lizzie cried [MASK] to sleep.")
print([r["token_str"] for r in result])
```

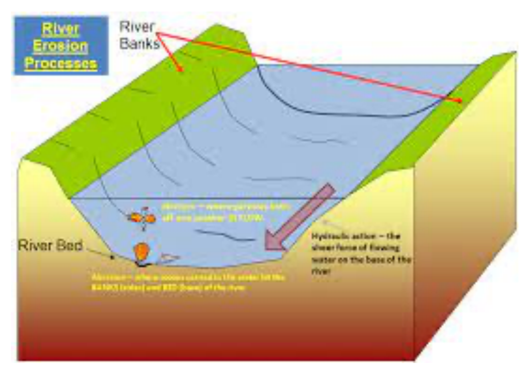
```
['themselves', 'herself', 'me', 'them', 'him']
```

- 

# LMs and word sense knowledge

- The word sense disambiguation (WSD) task is typically formulated as labeling words in context with their senses as defined by a dictionary or other lexical resource.

“The ***bank*** will not be accepting cash on Saturdays. ”

# LMs and word sense knowledge

- Many resources exist to support work on word senses.
- **WordNet**
  - Provides a fine-grained and comprehensive inventory of words and their senses for English.
- Several large annotated corpora have been constructed using WordNet senses,
  - SemCor
  - OntoNotes: has sense annotations for nouns and verbs,
  - Pattern Dictionary of English Prepositions (PDEP) corpus

# LMs and word sense knowledge

UNIVERSITÄT DUISBURG ESSEN

*Offen im Denken*

```
[58] result = unmasker("The bank will not be accepting cash on Saturdays. bank is a [MASK].")
     print([r["token_str"] for r in result])

     ['bank', 'failure', 'banks', 'mistake', 'business']

     result = unmasker("The river overflowed the bank. bank is a [MASK].")
     print([r["token_str"] for r in result])

     ['river', 'bank', 'lake', 'pond', 'wall']
```

# Summary

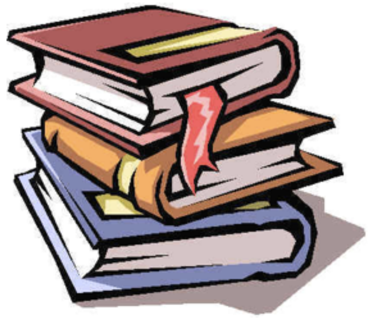
UNIVERSITÄT DUISBURG ESSEN
*Offen im Denken*

- **Pretrained language models (LMs)**
  - Unidirectional
  - Bidirectional
- **LMs as knowledge base**
  - LMs and factual knowledge
  - LMs and linguistic knowledge
  - LMs and word sense knowledge

# Readings

## Mandatory

- https://aclanthology.org/D19-1250.pdf
- https://arxiv.org/pdf/1901.05287.pdf
- https://aclanthology.org/2021.blackboxnlp-1.43.pdf



## Practice V

- Use your notebook in Google Colab (https://colab.research.google.com)
- Play with embeddings of some words
  - https://www.shanelynn.ie/word-embeddings-in-python-with-spacy-and-gensim/
  - **Check the relation between countries and cities**
  - **The word representation of which word is the nearest to the output vector of v(king) - v(man) + v(woman)?**
  - Relations between words in a language can be mapped to mathematical relations between their embeddings in an embedding space

# Practice VI

UNIVERSITÄT DUISBURG ESSEN

*Offen im Denken*

- Open GPT-3 playground: https://beta.openai.com/playground
- Give it some hints (a.k.a prompts) and let it complete the rest of the text?

  "This is a text about knowledge base systems. We aim at "
- Does it look knowledgeable?
- Test it for various properties of knowledge bases
  - *"Tail is part of a cat. Is this claim valid?"*
  - *"Birds can fly. is it correct?" Vs "Birds cannot fly. is it correct?"*
  - *"Musician is part of orchestra. Arm is par of a musician. Can we claim that arm is part of orchestra?"*
  - **Find an example that GPT-3 does not have any knowledge about?**