MUNI
FI

Deep Dive into Generative AI and Large Language Models

Mgr. Tomáš Foltýnek, Ph.D.

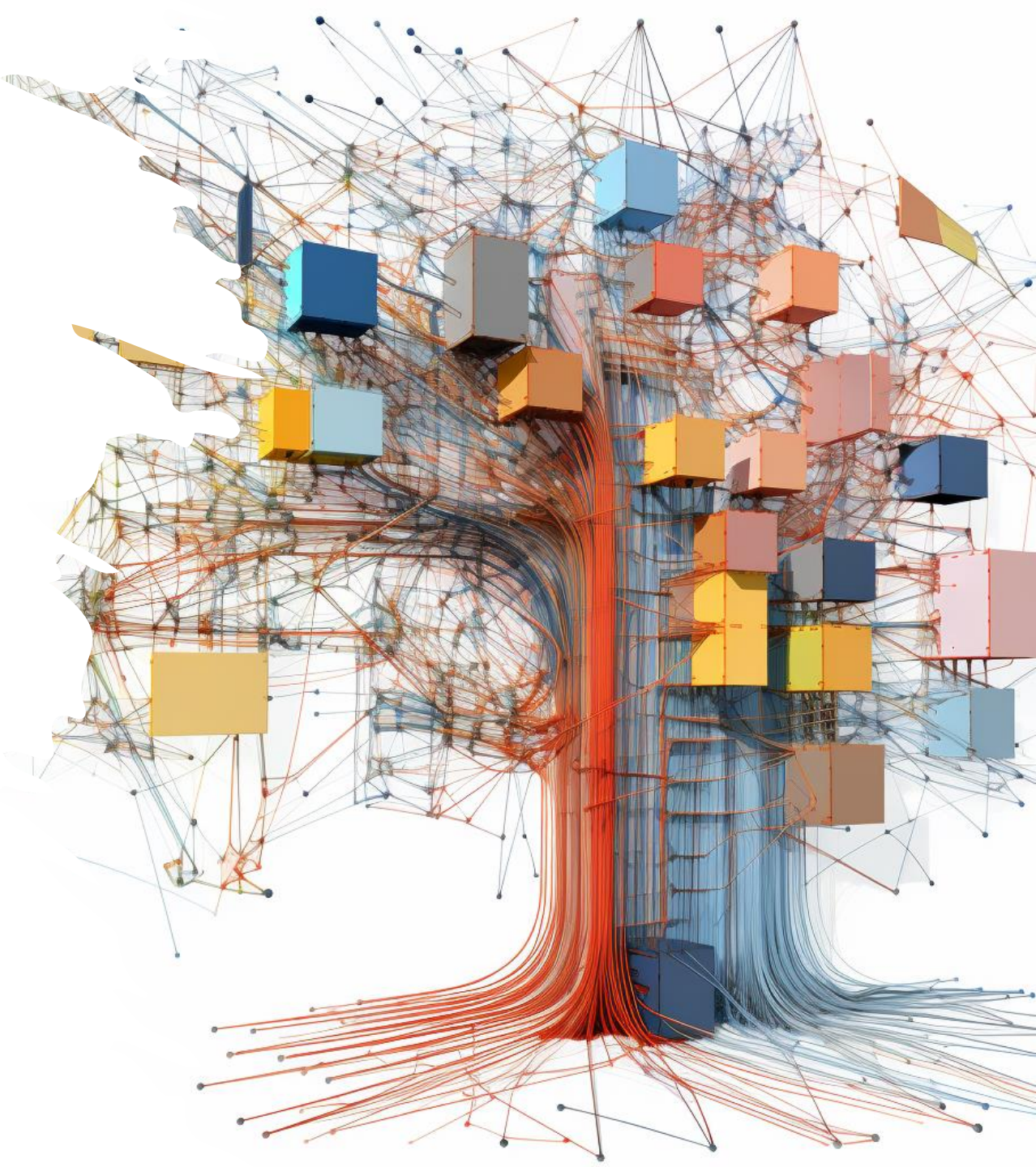
Department of machine learning and data processing

Faculty of Informatics, Masaryk University

foltynek@fi.muni.cz

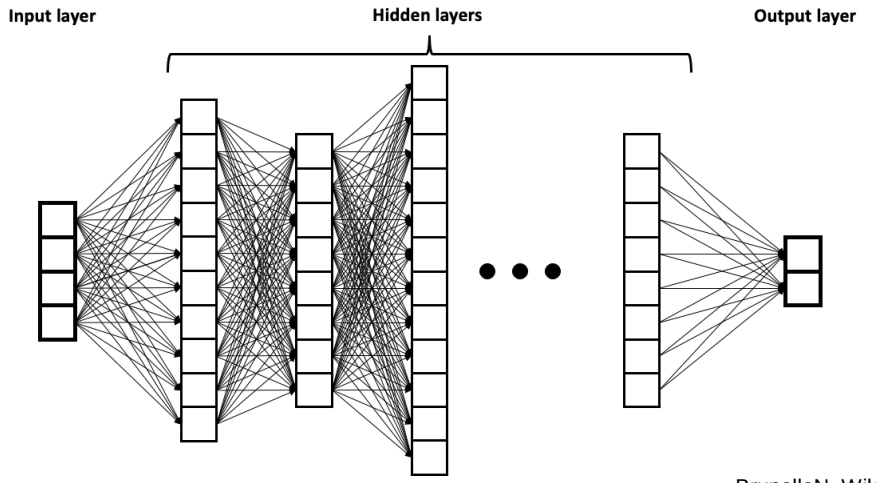
Image generated by Midjourney
/imagine artificial intelligence writing student essay in blue and yellow

How Generative Language Models Work



Machine Learning & Neural Networks

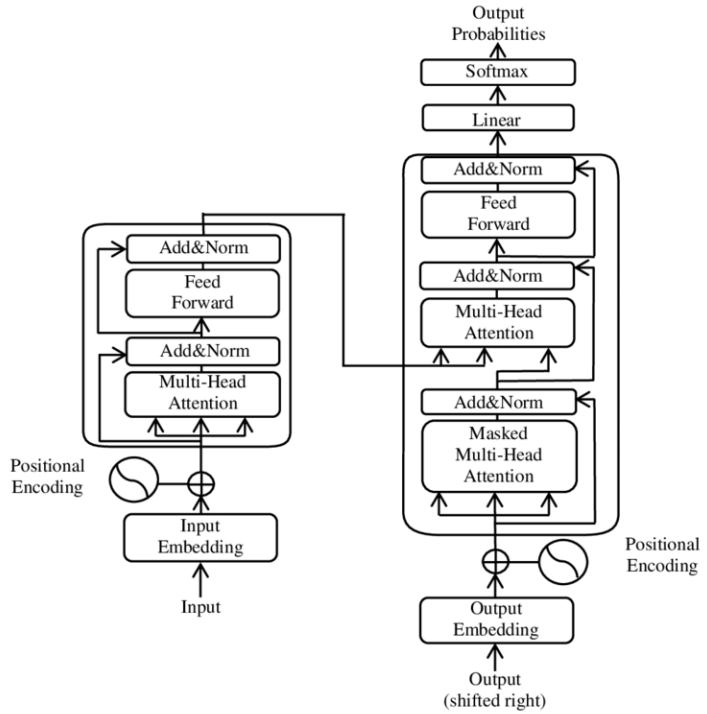
- Creation of single-purpose models (functions)
  - by analysing data
- Principles known since 1960s
- The programmer does not write the code, but defines the problem and specifies the criteria
- The model is built iteratively
- We talk about “training” or “learning”



BrunelloN, Wikimedia Commons

Transformers

- 2017 Google Brain (Attention is all you need)
- Encoding
  - Vector representation of each token
    - Token = basic unit (one word or part of word)
    - Based on word embeddings (i.e. context of words)
    - Attention (relations) between tokens
  - Feed-forward neural network
- Vector representation of the “meaning” of the input text
- Decoding
  - Based on the input from the encoder and the previous output of the decoder
  - Output vector → Output token
- Useful for many NLP tasks
  - Machine translation, paraphrase, summarization, question answering…

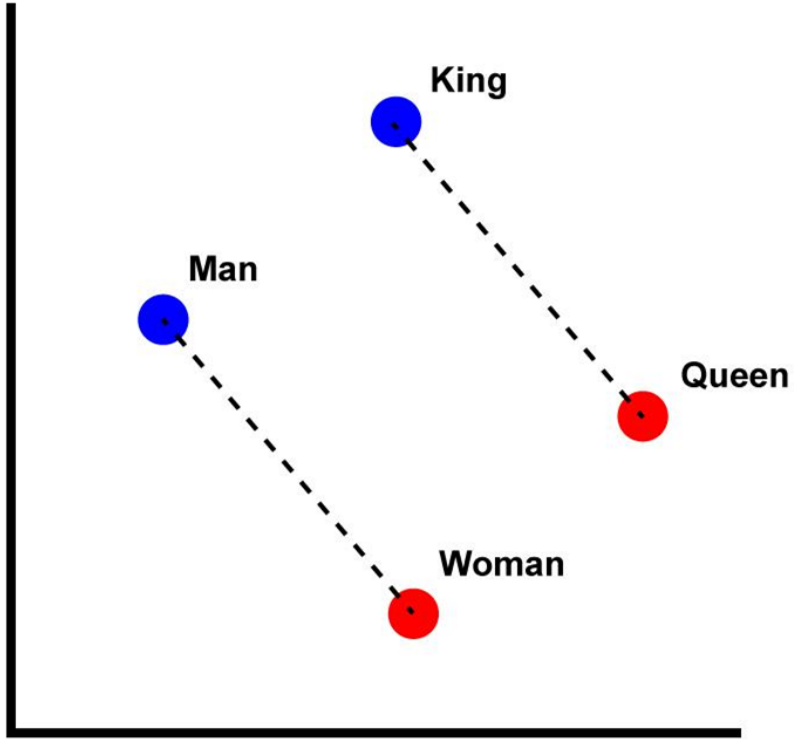


Latent Vector Space

- Images and texts are represented as **latent vectors**
- The position captures semantic relations among object

e.g. *King - Man + Woman = Queen*

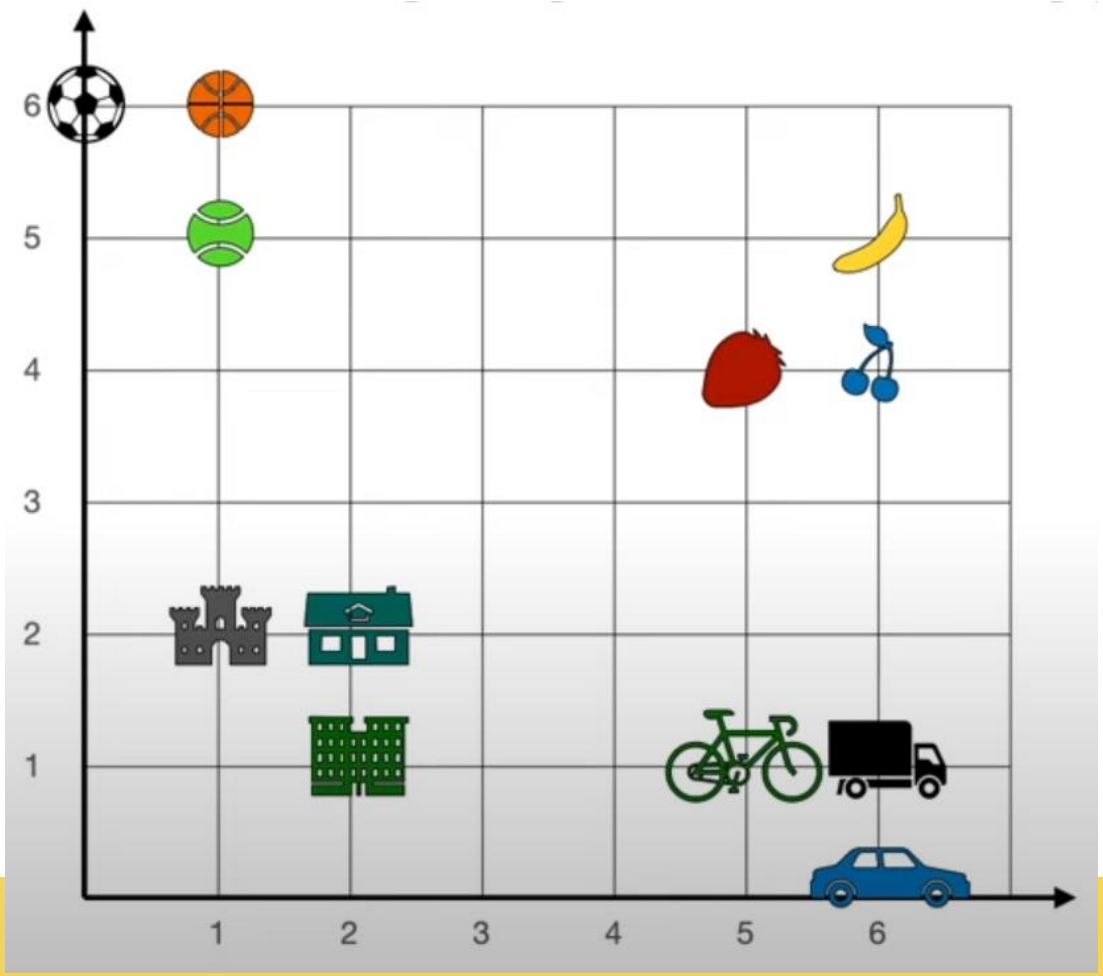
- Large language models (LLM)
  - multidimensional latent spaces



Word2Vec

Word Embeddings

Where would you put the word „apple“?



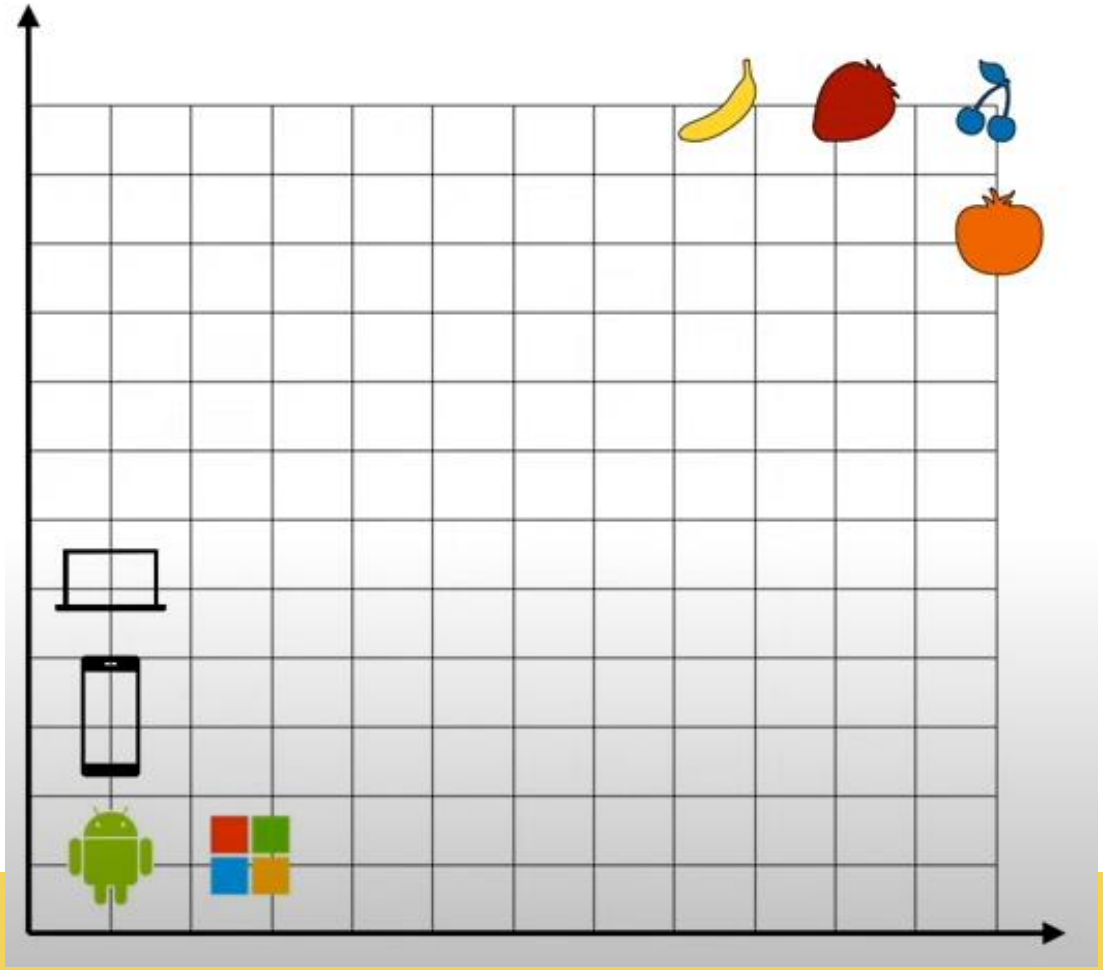
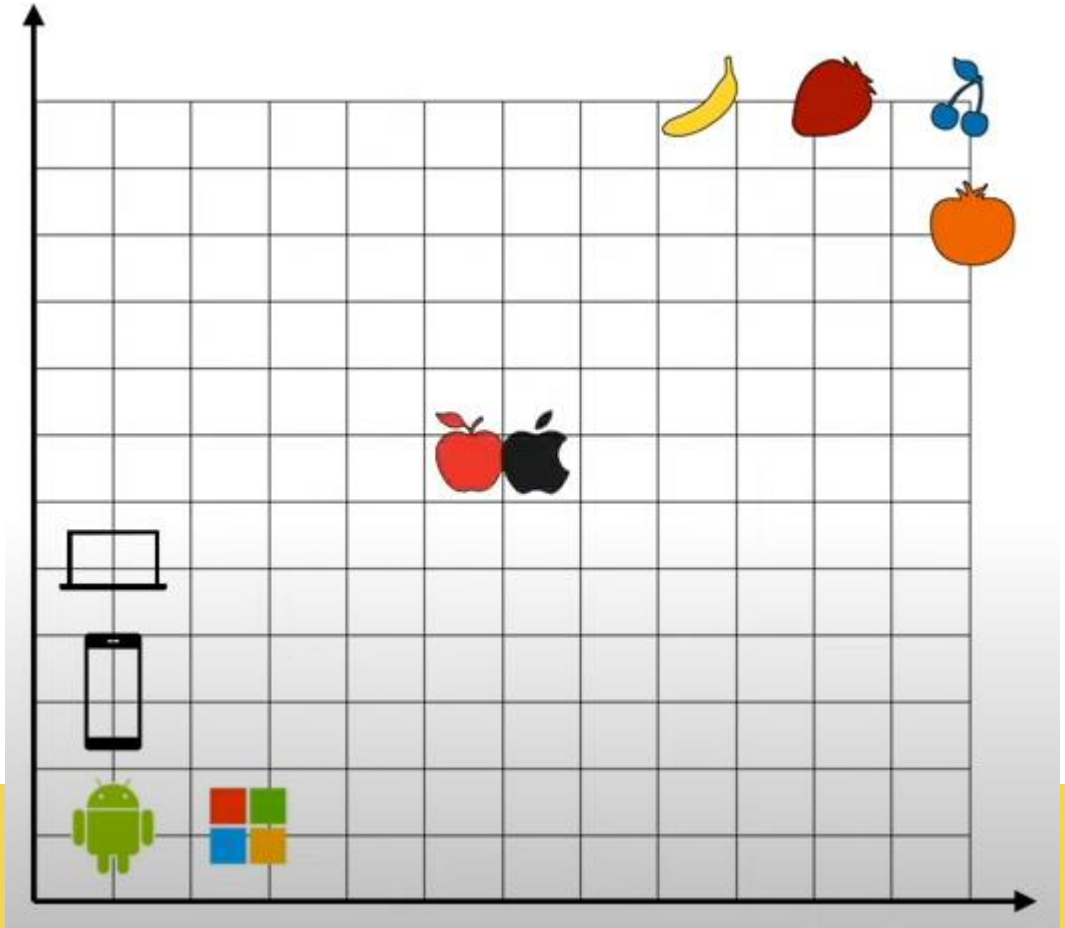


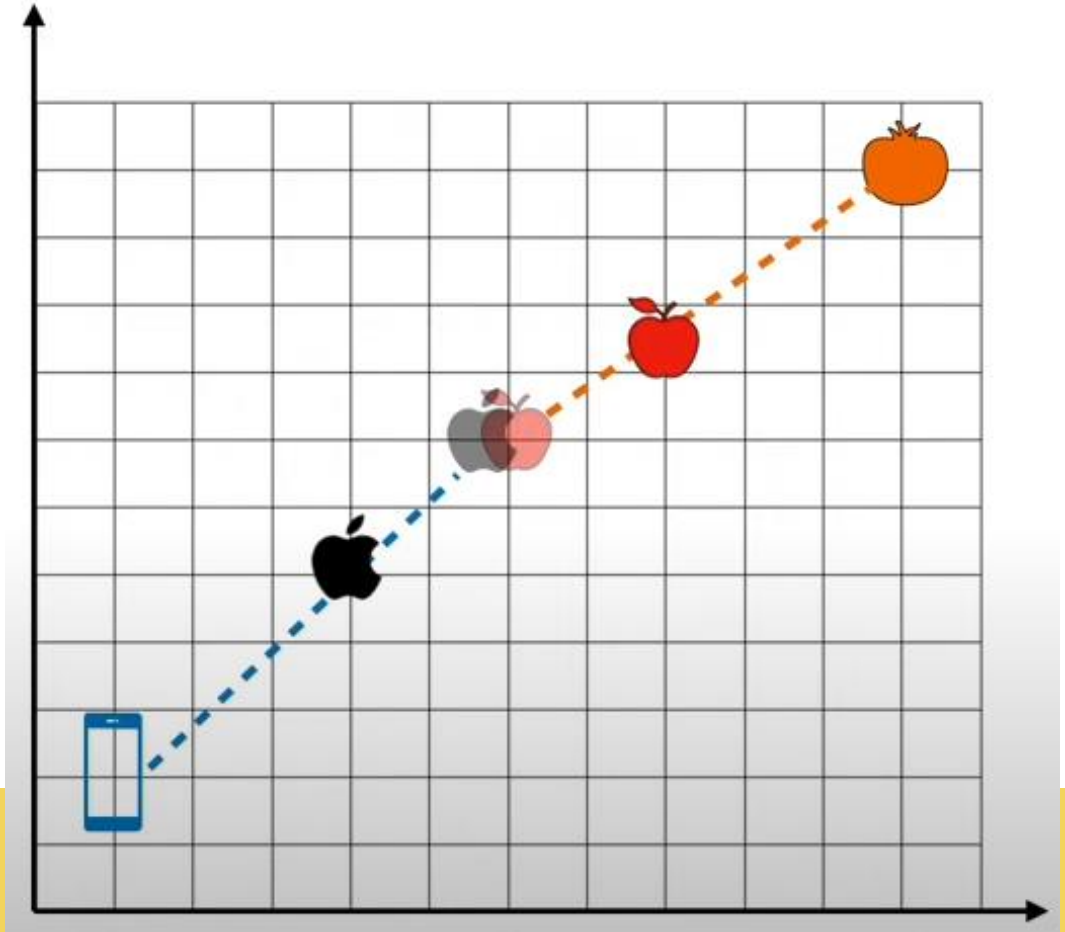
Image source: Serrano Academy

Image source: Serrano Academy

Attention

- I am going to eat an **apple** and an orange.
- **Apple** released a new model of iPhone.



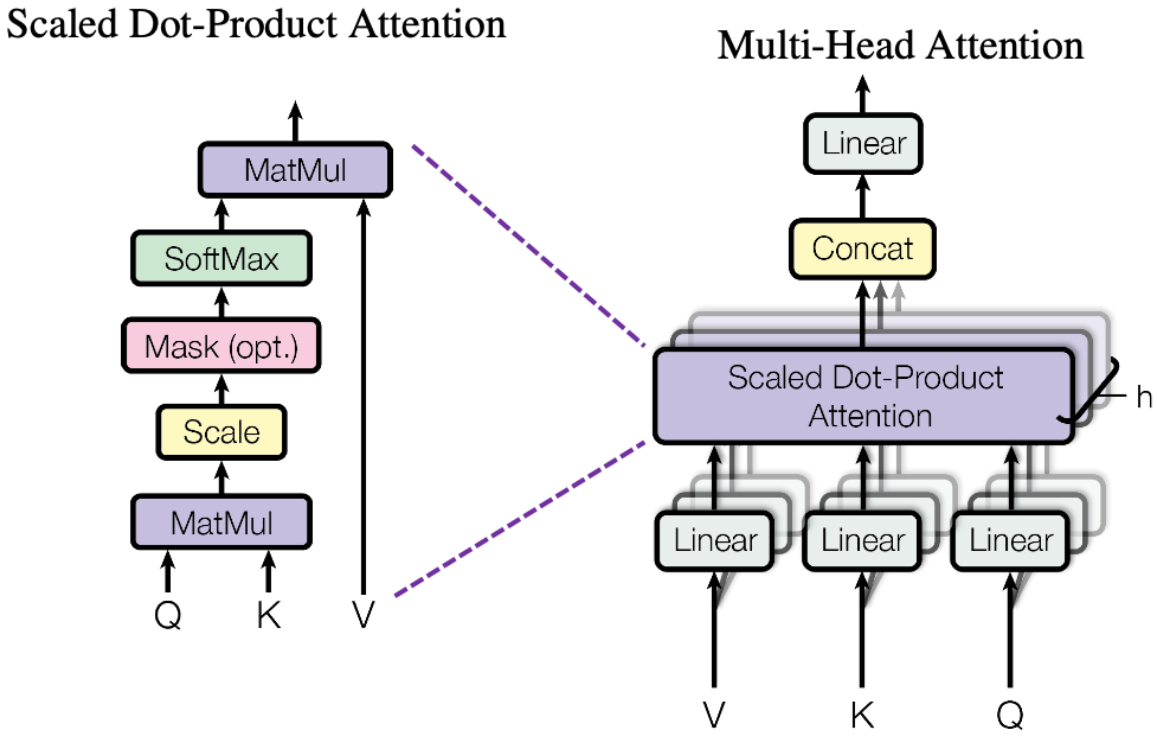


Context matters: (Self-)Attention

- Proximity pulls (like gravity)
- Compute attention matrix (proximity for each pair of words)
  - Simple dot product
  - Closer words ”pay attention” to each other
- Adjust the values of embeddings according to the matrix
  - Move the words in the vector space closer to those they attend to
  - Uses three pre-trained matrices: Keys, queries, values

$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V$$

Multi-Head Attention

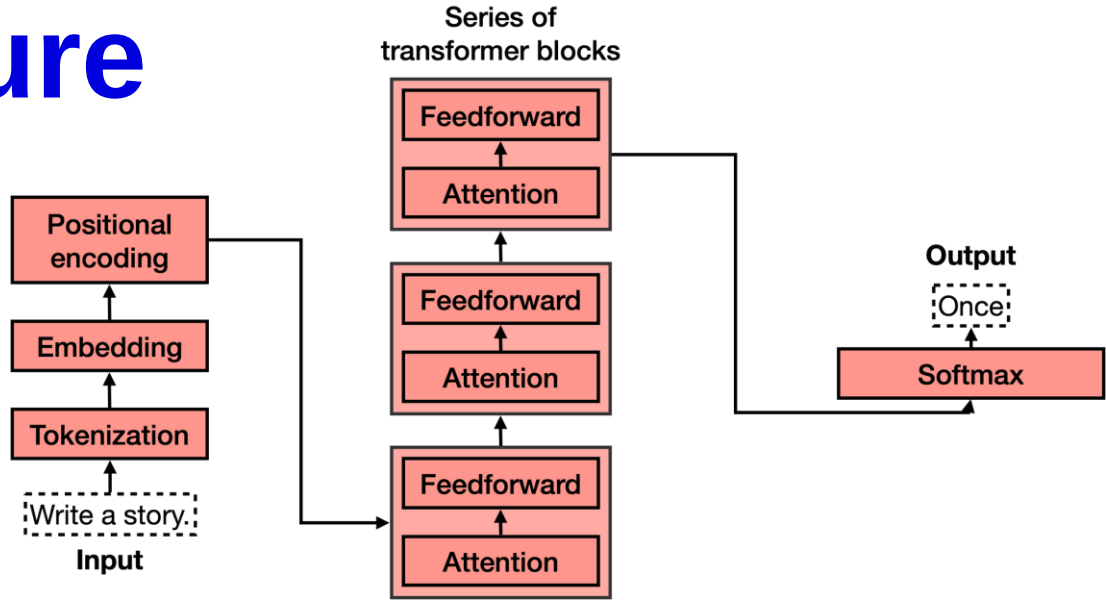


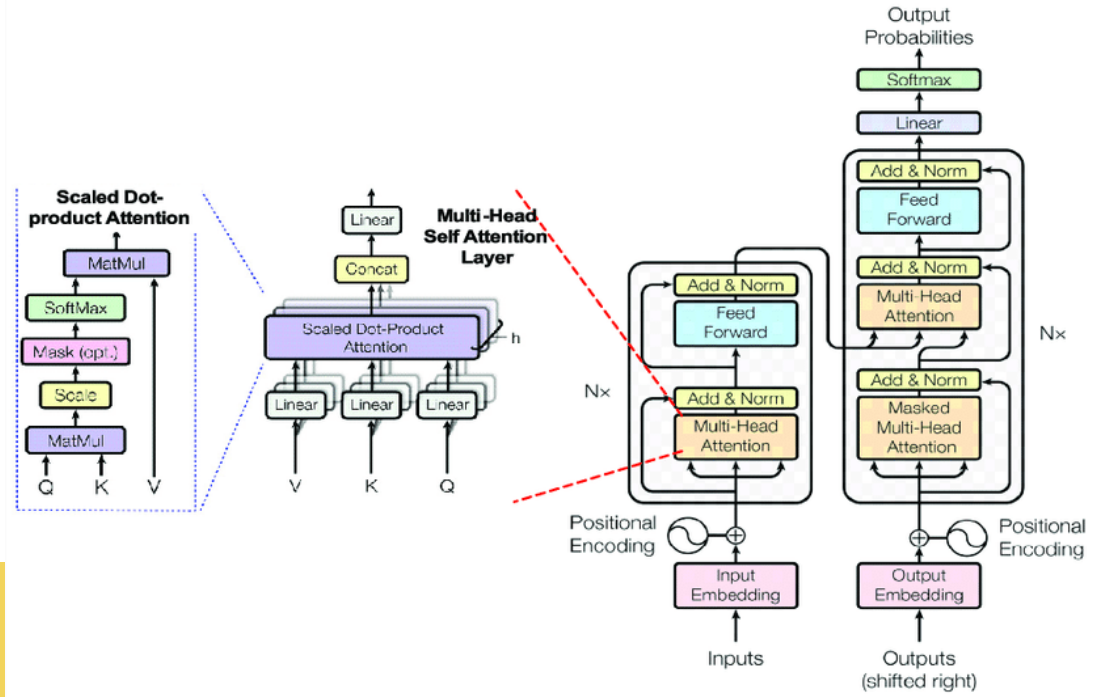
- One attention is not enough for more complex tasks
- We need to increase the model capacity
  - capture more features, e.g.
    - syntactic vs. semantic relations
    - genre, writing style
    - short-term vs. long-term dependencies
  - focus on different positions in the text
- Solution: Multi-head attention
  - The attention step is performed several times (in parallel)
  - The results are concatenated

Transformer Architecture

- Each block captures more features
- Higher-order congitive tasks require combination of the features
  - We need more blocks

- Autoregressive text generation
  - One token at a time
  - The output token becomes part of the input
  - The whole process repeats





Text Transformation vs. Text Generation

Text Transformation

- The LLM is fed with an already existing text
  - E.g. a research paper
- Task: Transform the text into another form
  - Translate, summarize, rephrase, extract information
- Lower risk of hallucination

Text Generation

- The LLM has no information but the prompt
- Task: Generate the text based on the prompt
  - Produces “statistical average” of all texts used for training
- Higher risk of hallucination

Ethical Challenges in Scientific Context

ENAI Working Group: Technology and Academic Integrity

- European Network for Academic Integrity
  - Association of HEIs and other institutions
  - Support and experience exchange in the field of academic integrity
    - Academic = research + education
  - Working group on Technology and Academic Integrity
- Wide availability of generative AI exacerbates current threats for academic integrity
  - ghostwriting, data fabrication and falsification,…
- May 2023: ENAI published a general recommendation
  - Editorial in the International Journal for Educational Integrity
- 2024: Research on ethical challenges of GenAI in science

Ethical Issues in Research Process

- Literature gathering
  - Unreliable sources (predatory journals)
- Textual understanding and summarization
  - Inaccurate or misleading information
- Code generation (data processing)
  - Wrong code, mistakes
- Data analysis
  - Inaccurate or misleading information
- Text creation
  - Hallucination, bias
- Communication of research results
  - Oversimplification, stereotypes
- All phases
  - IPR, data protection, privacy protection

beta

Does human activity influence Earth climate?

Synthesize

All years

Share This Page

beta These features are in Beta and not 100% accurate. Use with discretion. You can share feedback **here**

Summary

Consensus Meter

6 papers analyzed

- Yes - 100%
- Possibly - 0%
- No - 0%

The report judged it extremely likely that human activity is the predominant cause of recent climate change due to an increase in anthropogenic greenhouse gases in the atmosphere (IPCC, 2013).

• Yes

The Carbon Footprint as a Mediating Tool in Students' Online Reasoning about Climate Change

G. Fauville et al. | 2016

Cite | Share

(OPEN ACCESS ON ENCYCLOPEDIA WEBPAGE: http://climatescience.oxfordre.com/view/10.1093/acrefore/9780190228620.001.0001/acrefore-9780190228620-e-328) An overwhelming percentage of climate scientists agree that human activity is causing the global climate to change in ways that will have deleterious consequences both for the environment and for humankind.

• Yes

Climate Change Conspiracy Theories

J. Uscinski et al. | 2017

Cite | Share

Our results indicate that climatically induced ecological thresholds are not yet crossed, but that human activity has an increasing influence, particularly on the terrestrial ecosystem in our study area.

• Yes

Terrestrial and aquatic responses to climate change and human impact on the southeastern Tibetan Plateau during the past two centuries

beta

does vaccination cause autism?

Synthesize

All years

Share This Page

beta These features are in Beta and not 100% accurate. Use with discretion. You can share feedback **here**

Summary

Top 10 papers analyzed

These studies suggest that vaccinations are not associated with the development of autism or autism spectrum disorder, despite some caregivers' beliefs.

Consensus Meter

19 papers analyzed

- Yes - 5%
- Possibly - 26%
- No - 68%

Findings of this meta-analysis suggest that vaccinations are not associated with the development of autism or autism spectrum disorder.

• No

Vaccines are not associated with autism: an evidence-based meta-analysis of case-control and cohort studies.

Vaccine | Luke Taylor et al. | 2014

Meta Analysis | Highly Cited

Cite | Share

The authors concluded that The findings of this meta-analysis suggest that vaccinations are ot associated with the development of autism or autism spectrum isorder".

• Yes

Autism and vaccination: The value of the evidence base of a recent meta-analysis.

Vaccine | C. Turville et al. | 2015

Rigorous Journal | Meta Analysis

Cite | Share

The extensive research conducted so far indicates that childhood vaccination is not a cause of the sharp increase in autism diagnoses in recent decades.

• No

[Childhood vaccines and autism—much ado about nothing?].

Harefuah | I. Solt et al. | 2010

Cite | Share

Apologies for the oversight. To extract URLs from the metadata of a PD
`PdfFileReader` class from the PyPDF2 library to access the docume
Here's an updated version of the code:

```python
import PyPDF2
import re

def extract_urls_from_pdf(pdf_file_path):
    urls = []
    with open(pdf_file_path, 'rb') as pdf_file:
```

MP Try this again. I like your themes but your examples are not quotes from the database, they are fictional. you must create quotes Drawing only on the information presented in the abstracts of the papers.

I apologize for the confusion. Let me revise the themes and illustrative quotes, drawing only from the abstracts provided in the AIHE V1 database.

Theme 1: Transformative Potential of GenAI

Relevant Codes and Sub-Codes:

1. Benefits and Opportunities 1.1. Personalized Learning 1.2. Efficiency and Automation 1.5. Innovative Pedagogical Approaches
2. Applications and Use Cases 3.1. Writing Support and Feedback 3.3. Intelligent Grading and Assessment 3.6. Administrative Tasks and Support Services

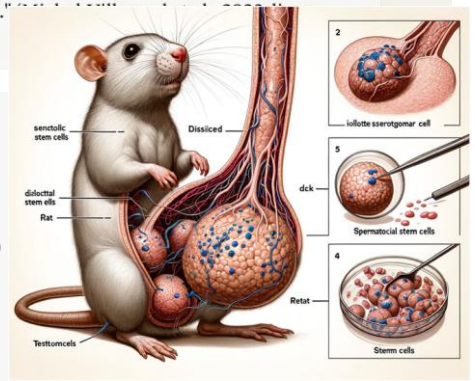
Illustrative Quotes:

1. "ChatGPT has the potential to revolutionize higher education by leveraging deep learning models to generate human-like content." 10-11)

Front. Cell Dev. Biol., 13 February 2024
Sec. Molecular and Cellular Reproduction
Volume 11 - 2023 | https://doi.org/10.3389/fcell.2023.1339390

RETRACTED: Cellular functions of spermatogonial stem cells in relation to JAK/STAT signaling pathway Retracted

Xinyu Guo[1] Liang Dong[2] Dingjun Hao[1]*



Foltýnek et al. International Journal of Educational Technology in Higher Education (2020) 17:46
https://doi.org/10.1186/s41239-020-00192-4

International Journal of Educational Technology in Higher Education

RESEARCH ARTICLE Open Access

Testing of support tools for plagiarism detection

Tomáš Foltýnek[1,2*], Dita Dlabolová[1], Alla Anohina-Naumeca[3], Salim Razı[4], Július Kravjar[5], Laima Kamzola[3], Jean Guerrero-Dib[6], Özgür Çelik[7] and Debora Weber-Wulff[8]

Who are the authors?

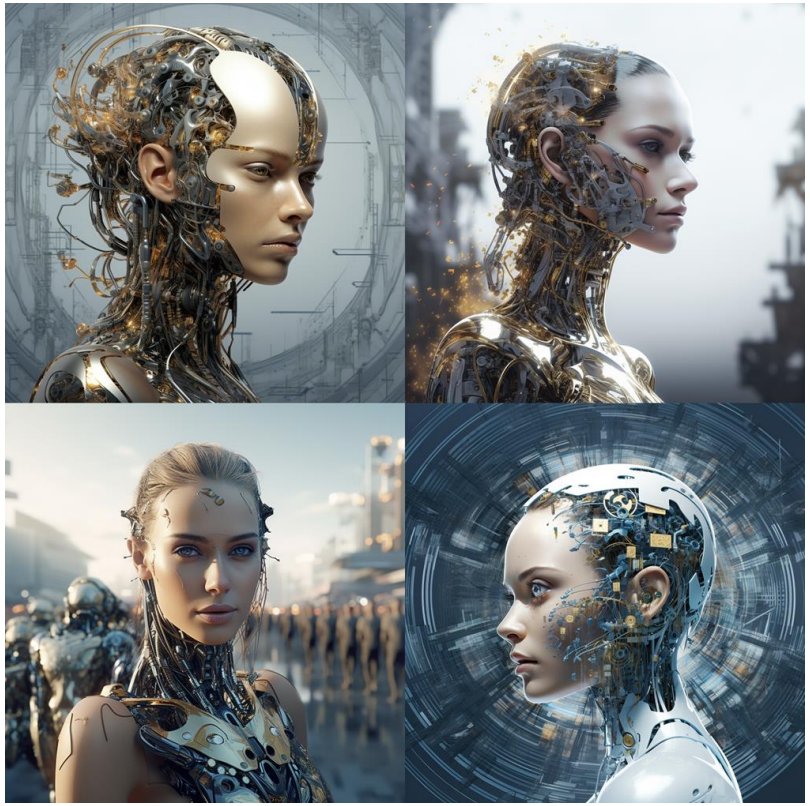
The authors of the study mentioned in the provided excerpts are:

1. Foltýnek, Tomáš: Department of Informatics, Faculty of Business and Economics, Mendel University in Brno, Czechia 30.
2. Budinska, Lenka: University of Wuppertal, Wuppertal, Germany 30.
3. Chalupová, Veronika: Riga Technical University, Riga, Latvia 30.
4. Can, Özden: Canakkale Onsekiz Mart University, Çanakkale, Turkey 30.
5. Dlabolová, Dita: Slovak Centre for Scientific and Technical Information, Bratislava, Slovakia 30.
6. Kremen, Vaclav: Universidad de Monterrey, Mexico, Mexico 30.
7. Kříž, Jan: Balikesir University, Balikesir, Turkey 30.
8. Wilhelm, Daniel W.: University of Applied Sciences HTW Berlin, Berlin, Germany 30.

These authors collaborated on the research project evaluating text-matching software for plagiarism detection, as detailed in the study 30.

Bias: Anecdotical Evidence

Midjourney: a professor, artificial intelligence, and a family meal



Solution?

- Thorough check and verification of outputs
- Proper testing of generated code
- Responsibility and accountability
- Raising awareness and education
  - IPR, data protection, privacy issues,…

- On the other hand:
  - Humans face the same ethical issues too
  - Inaccurate or misleading outputs, bias,…

MUNI
FI

Thank you!

Mgr. Tomáš Foltýnek, Ph.D.
foltynek@fi.muni.cz