

# Dekóderek

„Mi a következő token?”:  
A nyelvmodellezés elvei



# A nyelv és a nyelvmodell

- Nyelv: Tetszőleges véges  $A$  halmaz feletti  $A^*$  halmaz tetszőleges részhalmazát **nyelvnek** nevezzük, ahol  $A^*$  az  $A$  halmaz elemeiből alkotott füzérek halmaza. (Alberti 2006)
  - ♦ pl.  $A = \{a, b\}$ ,  $A^* = \{e, a, b, aa, ab, ba, bb, abc, \dots\}$
- Nyelvmodell: valószínűségi eloszlás  $A^*$  füzérei felett



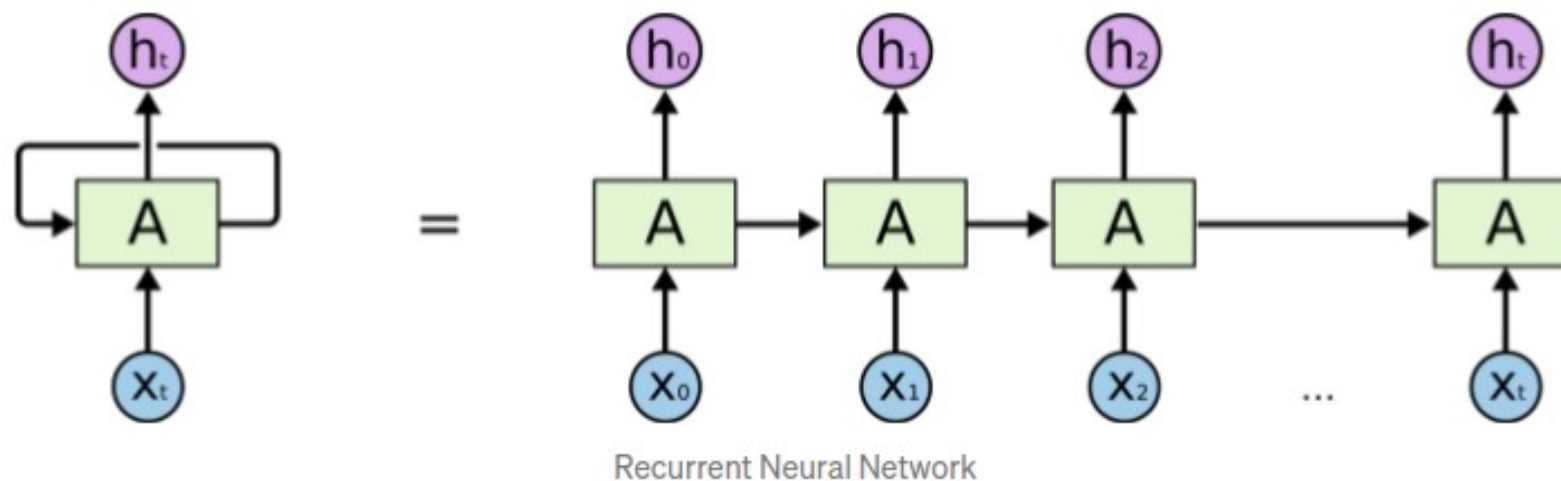
# Nyelvmodellezés feltételes valószínűséggel

- Legyen  $s$  egy füzér  $A$  halmaz elemeiből.
- Ekkor  $s = t_1 t_2 \dots t_n$ ,  $t_1, t_2, \dots, t_n \in A$
- $P(s) = P(t_1 t_2 \dots t_n) = P(t_1) P(t_2 | t_1) \dots P(t_n | t_1, t_2, \dots, t_{n-1})$



# Nyelvmodellezés RNN-nel

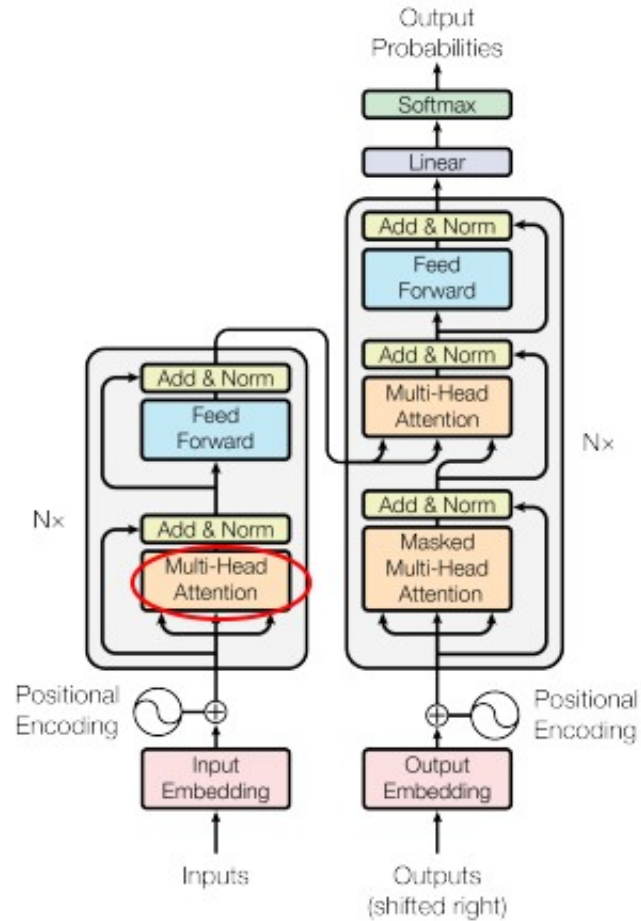
$$h_i = P(t_{i+1}|t_0, t_1, \dots, t_i)$$



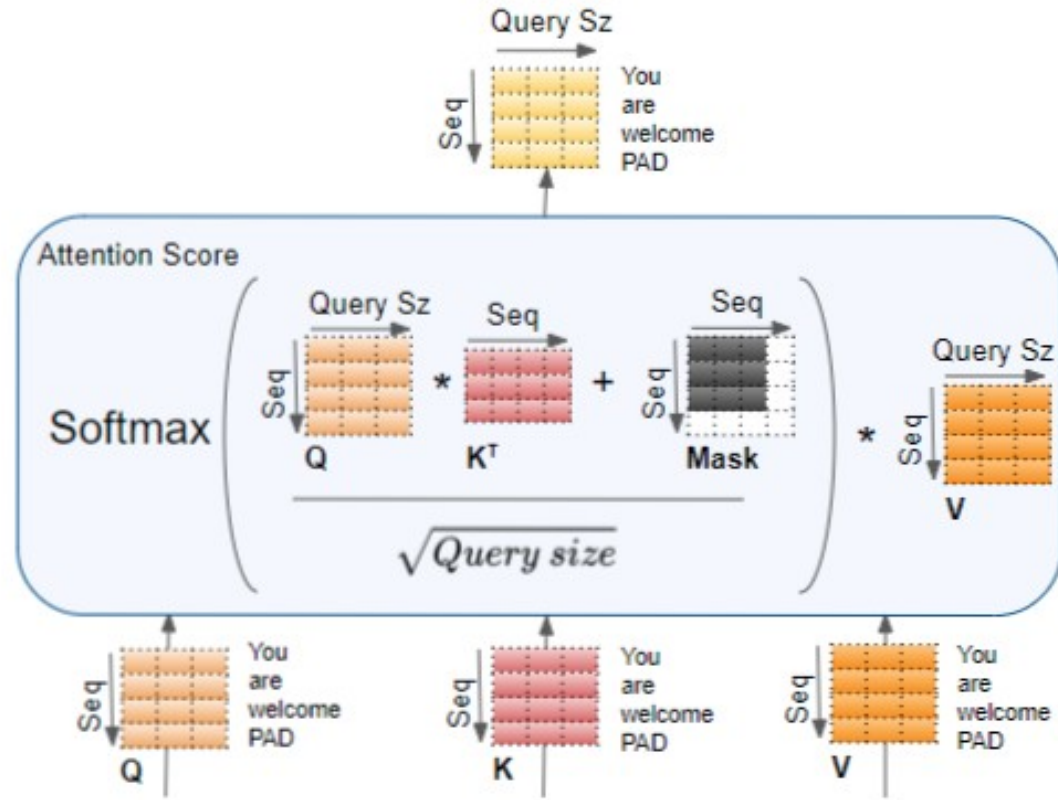
Forrás: Pranoy Radhakrishnan (2017): [Introduction to Recurrent Neural Network](#)



# Emlékeztető: a Transformer



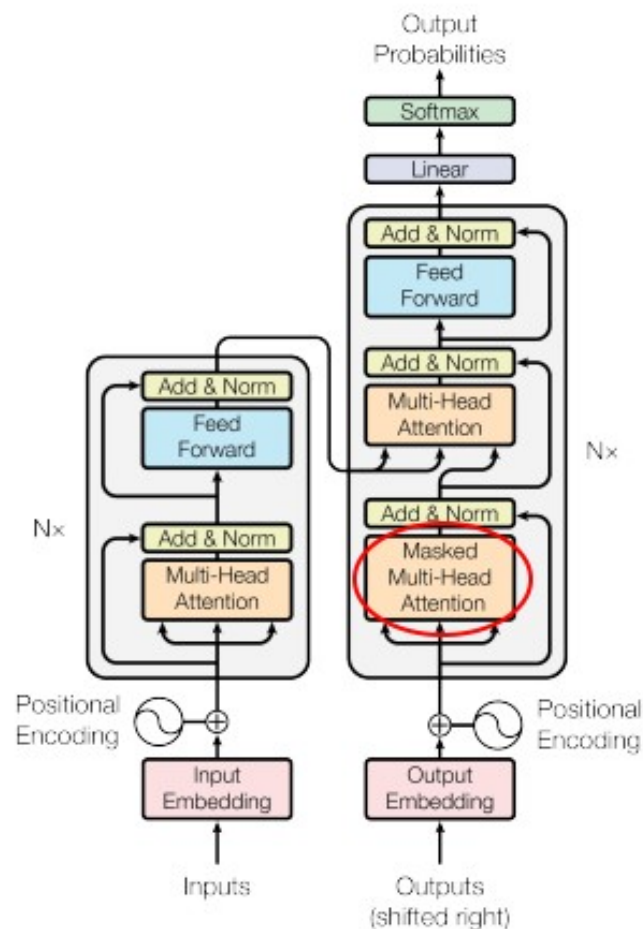
Eredeti kép forrása: Vaswani et al. (2017)



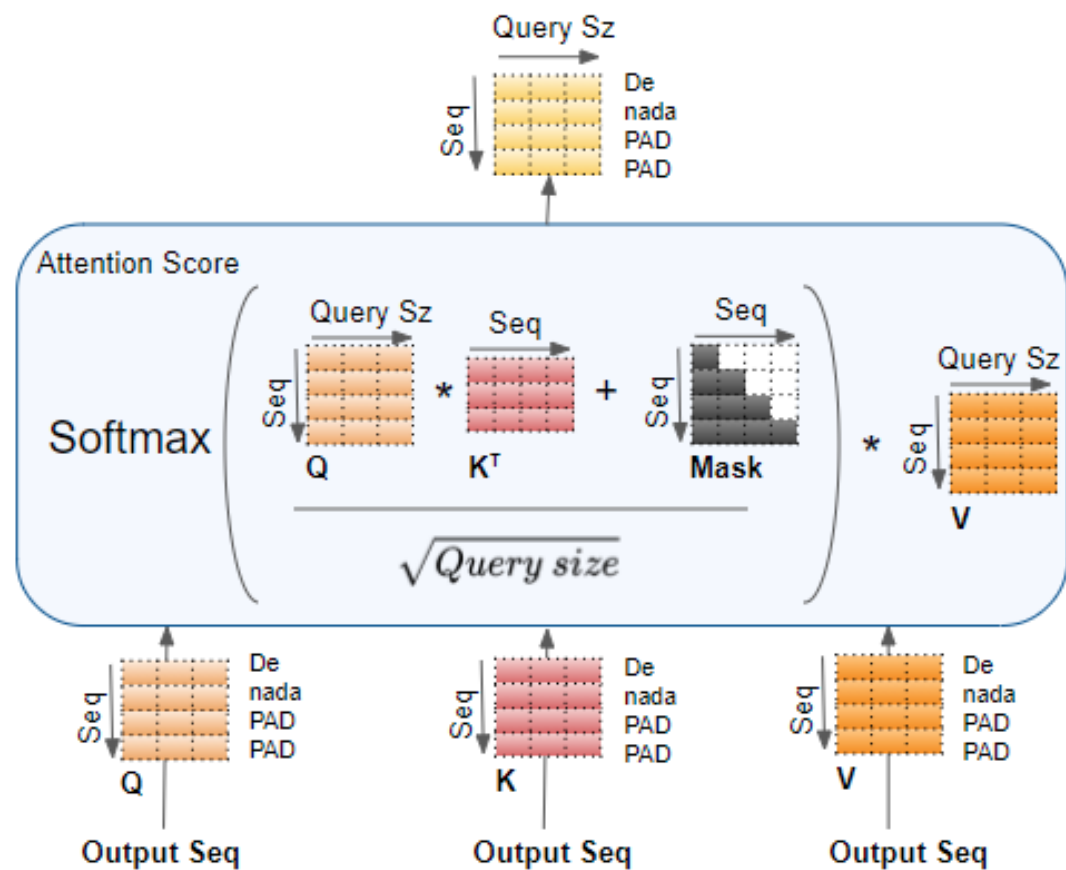
Forrás: Ketan Doshi (2021):  
Transformers Explained Visually (Part 3): Multi-head Attention, deep dive



# A Transformer dekóder



Eredeti kép forrása: Vaswani et al. (2017)

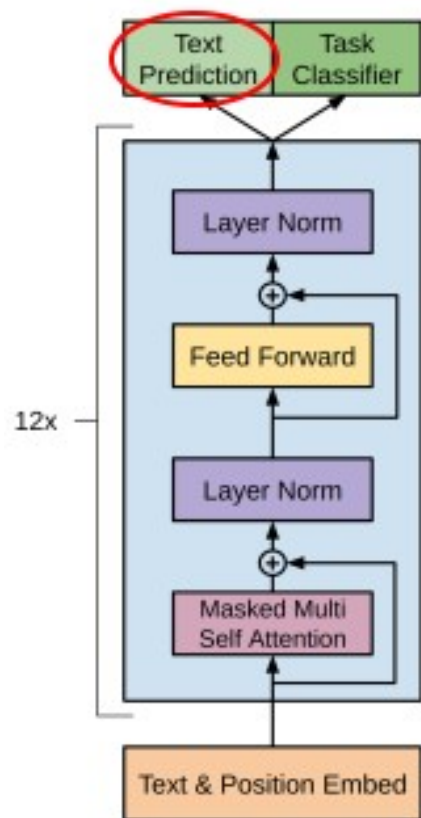


Forrás: Ketan Doshi (2021):  
[Transformers Explained Visually \(Part 3\): Multi-head Attention, deep dive](#)

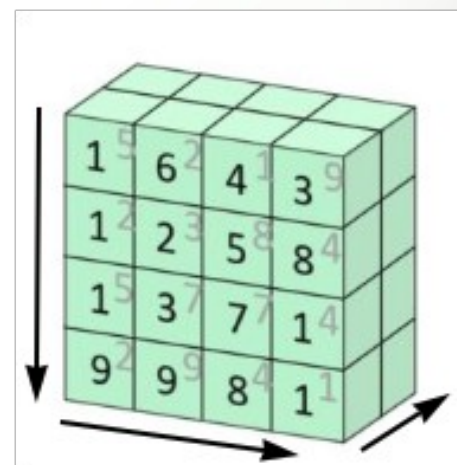




# A GPT



Eredeti kép forrása: Radford et al. (2018)



Eredeti kép forrása: [PyTorch – Tensors – The main data structure](#)



# Szöveggenerálás

- Mintavételezés valószínűségi eloszlásokból
- Algoritmusok: greedy, beam, top-k, top-p
- Top-k algoritmus: A valószínűséget újra elosztjuk, de csak a  $k$  legvalószínűbb token között, utána mintavételezünk





# Hogyan alkalmazható ez?

- Minden NLP feladat felfogható szöveggenerálásként
  - ◆ Kérdés – válasz: *Mi Spanyolország fővárosa? – Madrid.*
  - ◆ Osztályozás (szentiment): *Ma jó napom van! – 1*
  - ◆ Fordítás: *The cat is purring. – A macska dorombol.*
- Transfer learning: általános szöveggenerálás előtanítás alatt, specifikus szöveggenerálás finomhangolás alatt



# A transfer learning vége?

- Few-shot learning: Néhány példát megadunk kontextusként, majd folytatjuk a szöveget
- Zero-shot learning: Nincsenek példák, csak az előtanításra támaszkodik a modell



# GPT3 few-shot példa

```
Poor English input: I eated the purple berries.
Good English output: I ate the purple berries.
Poor English input: Thank you for picking me as your designer. I'd appreciate it.
Good English output: Thank you for choosing me as your designer. I appreciate it.
Poor English input: The mentioned changes have done. or I did the alteration that you
requested. or I changed things you wanted and did the modifications.
Good English output: The requested changes have been made. or I made the alteration that you
requested. or I changed things you wanted and made the modifications.
Poor English input: I'd be more than happy to work with you in another project.
Good English output: I'd be more than happy to work with you on another project.
```

---

```
Poor English input: Please provide me with a short brief of the design you're looking for and
that'd be nice if you could share some examples or project you did before.
Good English output: Please provide me with a brief description of the design you're
looking for and that would be nice if you could share some examples or projects you have
done before.
```

---

```
Poor English input: The patient was died.
Good English output: The patient died.
```

---

```
Poor English input: We think that Leslie likes ourselves.
Good English output: We think that Leslie likes us.
```

---

```
Poor English input: Janet broke Bill on the finger.
Good English output: Janet broke Bill's finger.
```

---

```
Poor English input: Mary arranged for, in St. Louis, John to rent a house cheap.
Good English output: Mary arranged for John to rent a house in St. Louis.
```

---

```
Poor English input: Today I have went to the store to to buys some many bottle of water.
Good English output: Today I went to the store to buy some bottles of water.
```

---

```
Poor English input: I have tried to hit ball with bat, but my swing is has miss.
Good English output: I tried to hit the ball with the bat, but my swing missed.
```

Forrás: Brown et al. (2020)



# FLAN zero-shot példa

## Input:

Movie review: This movie is the best RomCom since Pretty Woman.

Did this critic like the movie?

OPTIONS

-yes  
-no

## FLAN output:

yes

## Input:

Movie review: This movie is the best RomCom since Pretty Woman.

Did this critic dislike the movie?

OPTIONS

-yes  
-no

## FLAN output:

no

Forrás: Wei et al. (2021)



# Cikkek, linkek

Alberti Gábor (2006): *Matematika a természetes nyelvek leírásában*. Budapest: Tinta kiadó.

Brown et al. (2020): *Language Models are Few-Shot Learners*. [Link](#)

Radford et al. (2018): *Improving Language Understanding by Generative Pre-Training*. [Link](#)

Radford et al. (2019): *Language Models are Unsupervised Multitask Learners*. [Link](#)

Vaswani et al. (2017): *Attention is all you need*. [Link](#)

Wei et al. (2021): *Finetuned Language Models Are Zero-Shot Learners*. [Link](#)

Ketan Doshi (2021): *Transformers Explained Visually (Part 3): Multi-head Attention, deep dive*. (Towards Data Science) [Link](#)

Patrick von Platen: *How to generate text: using different decoding methods for language generation with Transformers*. [Link](#)

Pranoy Radhakrishnan (2017): *Introduction to Recurrent Neural Network*. (Towards Data Science) [Link](#)

*PyTorch – Tensors – The main data structure*. (Data Hacker) [Link](#)