

# (G)AI DML LLM

Štěpán Beneš



net  
via







**Máme se zajímat o LLM ?**

Myslivec jde z hospody domů do hájenky. Jeho psík se už těší domů za paní myslivcovou a tak běží domů napřed. Když doběhne domů, zasteskne se mu po pánovi a běží mu zase v ústrety. A tak stále dokola, dokud myslivec nedojde domů do hájenky. Kolik naběhá pes, pokud běhá rychlostí 20km/h, myslivec jde rychlostí 5km/h a z hospody do hájenky je to 10km.





Myslivec jde z hospody rychlostí 5 kilometrů za hodinu, domů do hájenky to má 10 kilometrů. Má psa, který běhá rychlostí 20 kilometrů neustále mezi ním a hospodou. Kolik naběhá pes, než myslivec dojde z hospody domů ?

Zobrazit koncepty



### Řešení

Nejprve zanalyzujeme pohyb psa a myslivce:

#### Myslivec:

- Vzdálenost: 10 km
- Rychlost: 5 km/h

#### Pes:

- Rychlost: 20 km/h
- Běhá tam a zpět mezi myslivcem a hospodou

#### Výpočet:

##### 1. Cesta psa za jednu cestu:

Pes běží z hospody k myslivcovi a pak zpět, takže urazí vzdálenost  $2 \cdot 10 \text{ km} = 20 \text{ km}$ .

##### 2. Počet cest psa:

Myslivec ujde 10 km rychlostí 5 km/h, takže mu to bude trvat  $10 \text{ km} / 5 \text{ km/h} = 2$  hodiny.

Během 2 hodin pes uběhne  $20 \text{ km/h} \cdot 2 \text{ h} = 40 \text{ km}$ .

#### Odpověď:

Pes naběhá **40 kilometrů**, než myslivec dojde z hospody domů.

#### Poznámka:

V tomto řešení předpokládáme, že pes běhá nepřetržitě rychlostí 20 km/h, dokud se myslivec nedostane domů. Ve skutečnosti pes pravděpodobně bude dělat pauzy, takže skutečná vzdálenost, kterou uběhne, bude o něco menší.



You

Myslivec jde z hospody rychlostí 5 kilometrů za hodinu, domů do hájenky to má 10 kilometrů. Má psa, který běhá rychlostí 20 kilometrů neustále mezi ním a hospodou. Kolik naběhá pes, než myslivec dojde z hospody domů ?



ChatGPT

Abychom zjistili, kolik kilometrů naběhá pes během cesty myslivce z hospody domů, musíme zvážit dobu, kterou myslivec stráví cestou.

Myslivec jde rychlostí 5 kilometrů za hodinu a má dojít 10 kilometrů. Z toho vyplývá, že mu to zabere:

$$\text{Čas} = \frac{\text{Vzdálenost}}{\text{Rychlost}} = \frac{10 \text{ km}}{5 \text{ km/h}} = 2 \text{ hodiny}$$

Pes běhá rychlostí 20 kilometrů za hodinu. Během dvou hodin, co myslivec putuje, pes tedy naběhá:

$$\text{Vzdálenost, kterou naběhá pes} = \text{Rychlost psa} \times \text{Čas} = 20 \text{ km/h} \times 2 \text{ h} = 40 \text{ km}$$

Takže pes naběhá 40 kilometrů, než myslivec dojde domů.



# OSNOVA

- Vysvětlení pojmů
- Neuronové sítě
- Jazykové modely
- Základní tipy pro práci s LLM

# Umělá inteligence

Umělá inteligence (AI) je obor informatiky, který se zabývá tvorbou inteligentních systémů, které dokáží:

- Vnímat své prostředí
- Uvažovat
- Učit se
- Řešit problémy autonomně

Cílem AI je vytvořit stroje, které dokáží vykonávat úkoly, které obvykle vyžadují lidskou inteligenci, jako je hraní her, rozpoznávání obrázků, překlad jazyků a psaní textu.

Generální umělá inteligence (GAI), je hypotetický typ AI, který by měl obecnou inteligenci srovnatelnou s lidskou.

# Procedurální řešení

## IF ... THEN ... ELSE

- Algoritmizace
- Program není nikdy “chytřejší” než tvůrce
- Složitě řešitelné úlohy:
  - Rozpoznání obrázku
  - Hry - šachy, GO
  - Pohybová mechanika
- Predikovatelné, snadno laditelné, opravitelné

```
31 def __init__(self, settings):
32     self.file = None
33     self.fingerprints = self()
34     self.logdups = True
35     self.debug = debug
36     self.logger = logging.getLogger(__name__)
37     if path:
38         self.file = open(os.path.join(path, 'requests.txt'), 'w')
39         self.file.seek(0)
40         self.fingerprints.update({request: 0} for request in self())
41
42 @classmethod
43 def from_settings(cls, settings):
44     debug = settings.getbool('debug', False)
45     return cls(job_dir(settings), debug)
46
47 def request_seen(self, request):
48     fp = self.request_fingerprint(request)
49     if fp in self.fingerprints:
50         return True
51     self.fingerprints.add(fp)
52     if self.file:
53         self.file.write(fp + os.linesep)
54
55 def request_fingerprint(self, request):
56     return request_fingerprint(request)
```

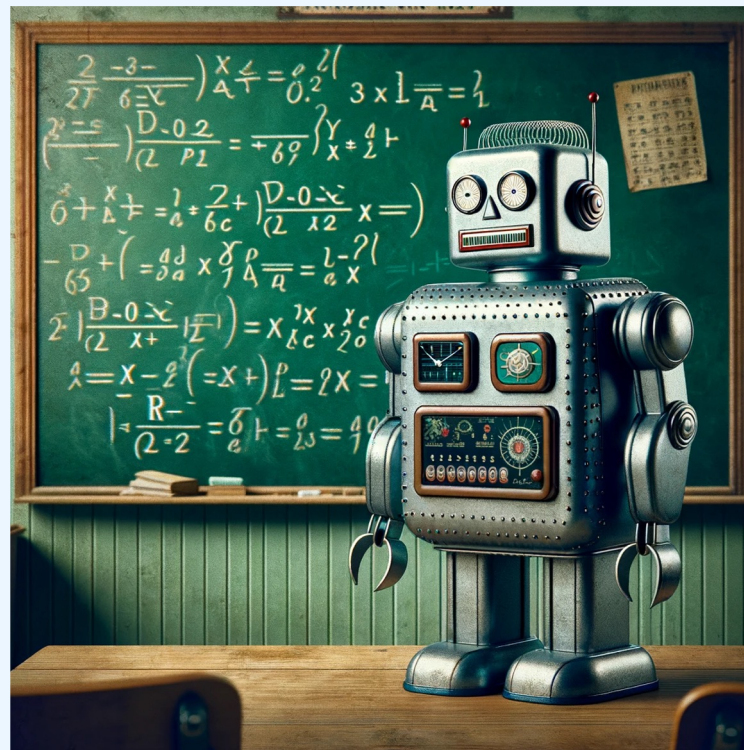


# Strojové učení

Metoda implementace AI.

Aplikace učící se z dat a zlepšují svůj výkon bez explicitního programování.

- **Neuronové sítě**
- Support Vector Machines - SVM
- Naivní Bayesův klasifikátor
- Rozhodovací stromy
- ...



# Historie

## 1951 Stochastic Neural Analog Reinforcement Calculator (SNARC)

- 40 neuronů
- “Kryší” problém - únik z bludiště

## 1957 Mark I Perceptron

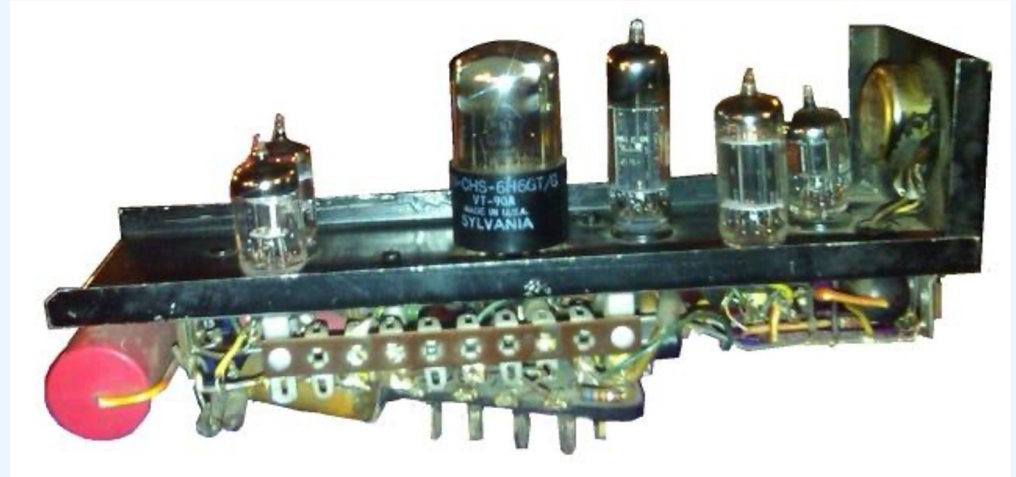
- Rozpoznávání znaků 20x20
- Váhy pomocí potenciometrů

## 198X rozvoj neuronových sítí

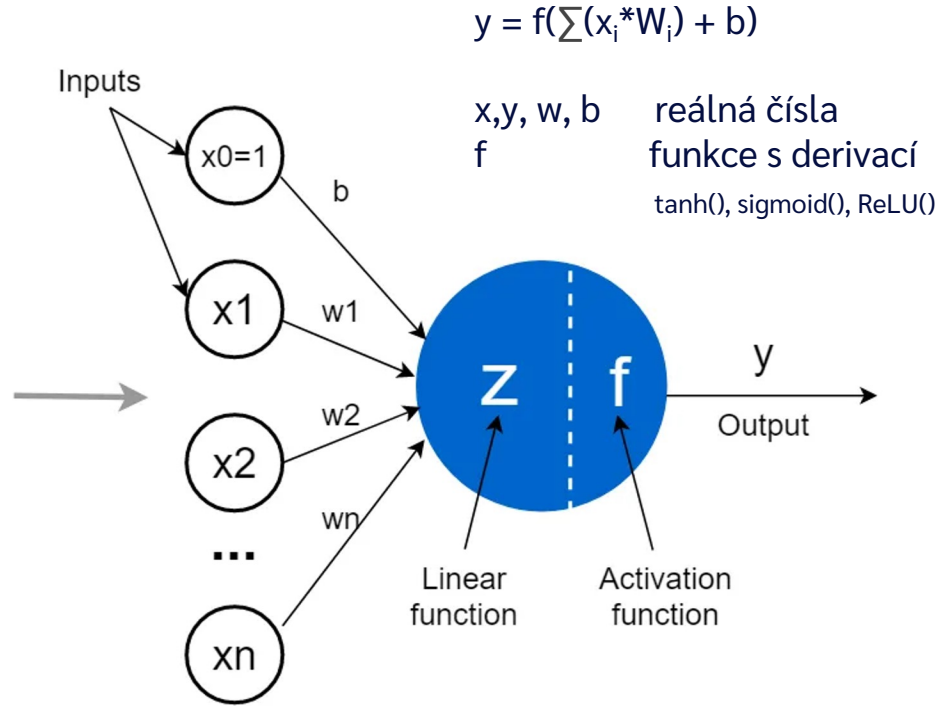
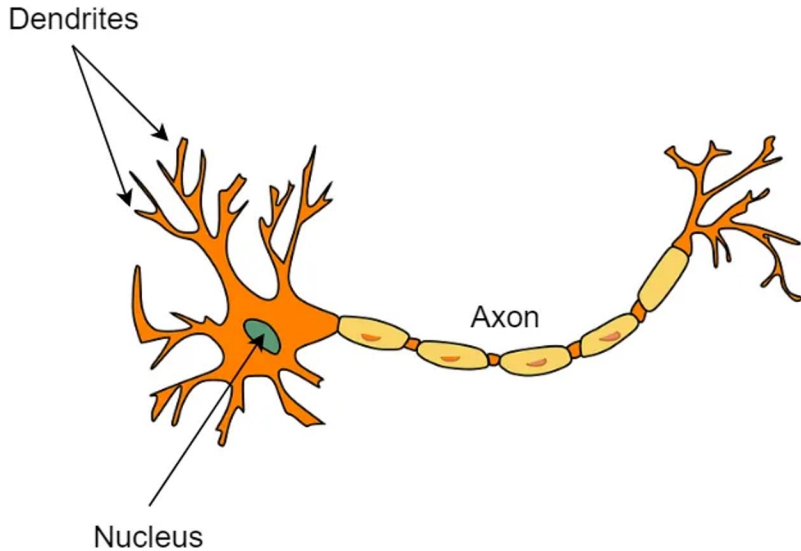
## 199X jiné technologie (SVM)

## 2005+ renesance neuronových sítí

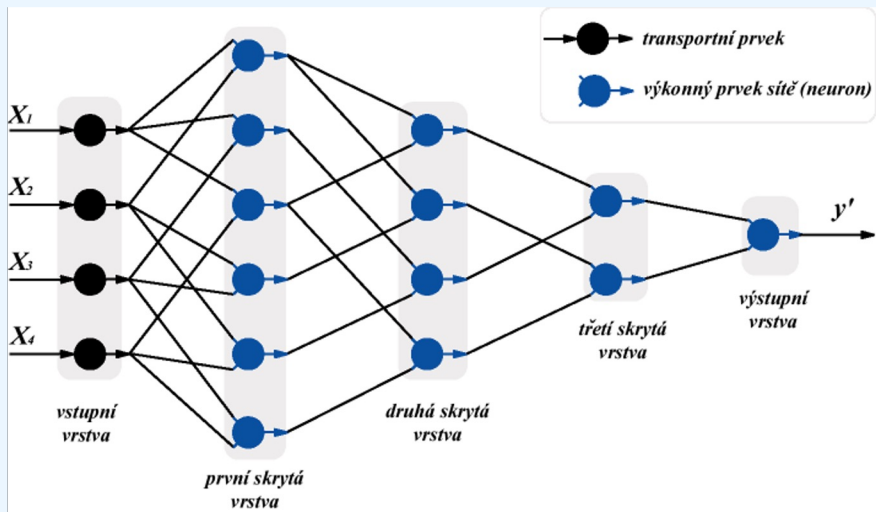
- **Deep** machine learning
- Výpočetní výkon, GPU
- SW implementace



# NEURON



# NEURONOVÁ SÍŤ



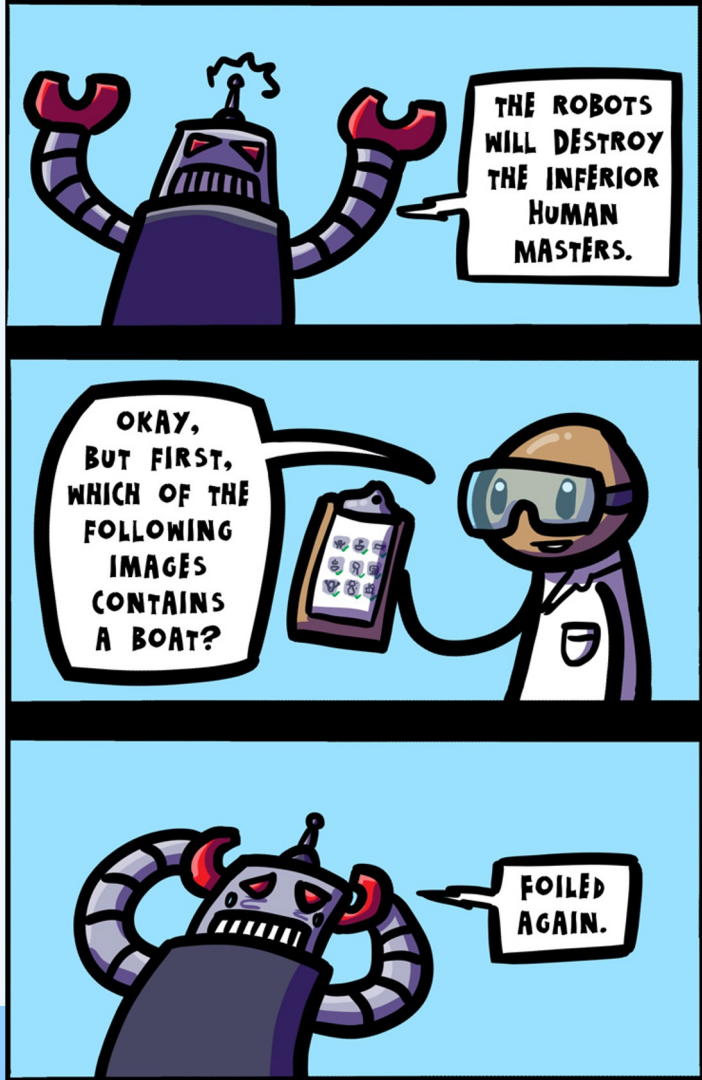
# Trénování sítě

- BackError Propagation
- Učení s dohledem
  - Data -> správná odpověď
  - Např. klasifikace obrázků
  - Limitem je učitel
- Učení bez dohledu
  - Korelace dat
  - Zajímavé vlastnosti
- Učení s posilováním
  - Reinforcement Learning
  - Akce -> odměna
  - Šachy, Go
  - Řízení auta



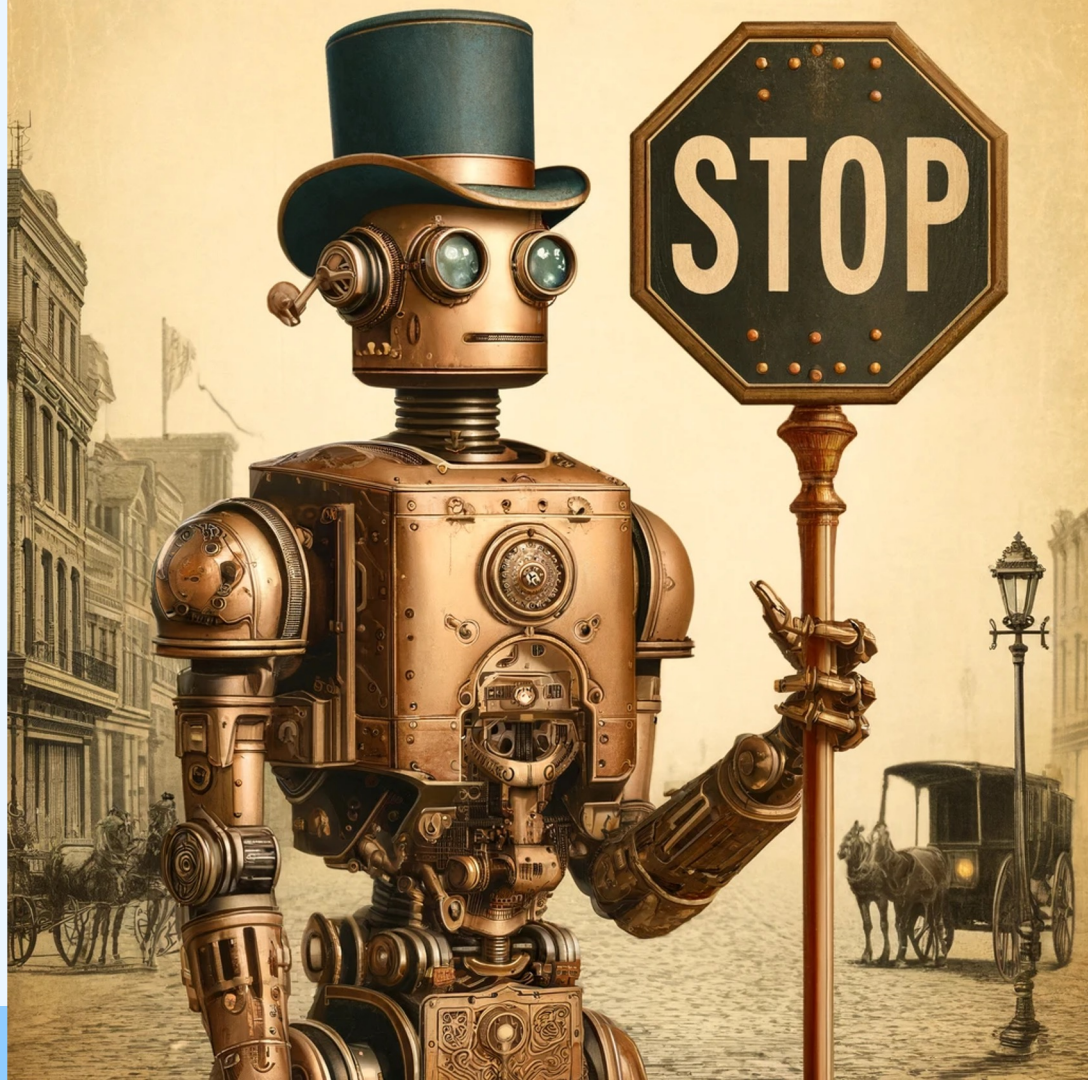
# Aplikace

- V posledních 10 letech obrovský rozvoj
- Rozpoznávání řeči, textu, obrázků
- Anomálie, detekce stavů
- Vyhledávání korelace dat
- ...
- LLM



# Problémy ML

- “Blackbox”
- Závislost na kvalitě dat
- Obtížnost ladění a oprav
- Přístup k datům

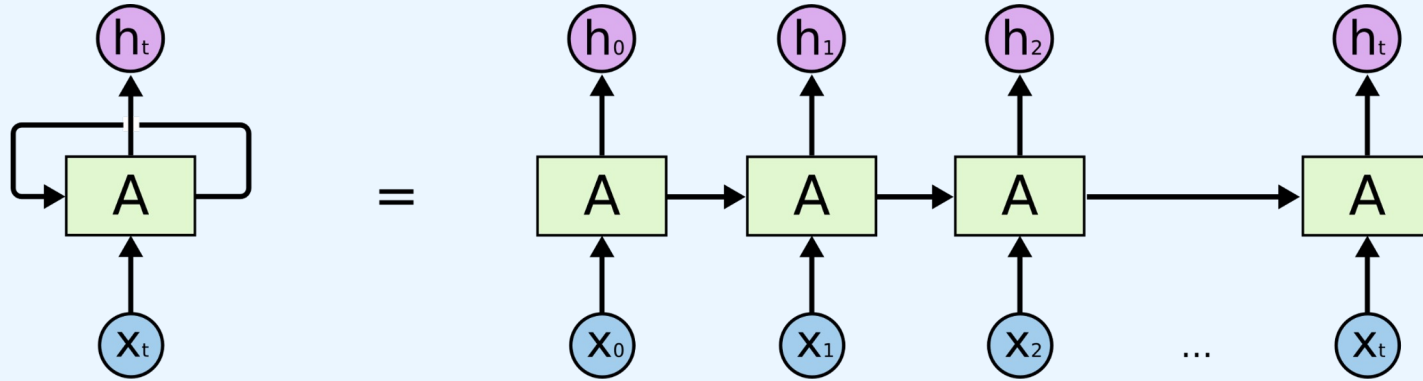


The background of the image is a dark, starry space scene with nebulae and distant galaxies. Overlaid on this is a complex, glowing network of white and blue lines connecting numerous small, bright nodes. The network is centered around a bright, multi-colored (yellow, orange, and white) light source that radiates outwards. The overall aesthetic is high-tech and digital, suggesting a global or data network.

**LLM - Large language model**

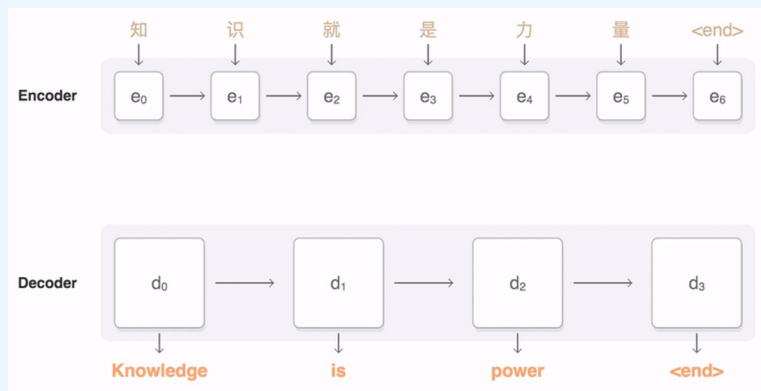


# Rekurentní neuronové sítě (RNN)



- “Paměť”
- K čemu se hodí ?
  - Zpracování dat s kontextem
  - Aplikace
    - Zpracování textu
    - Redukce šumu
    - Hledání datových korelací

# Seq2Seq



- Sekvenční zpracování - změna jedné sekvence na jinou
- Google (translator) vs. Tomáš Mikulov
- Encoder <-> Decoder

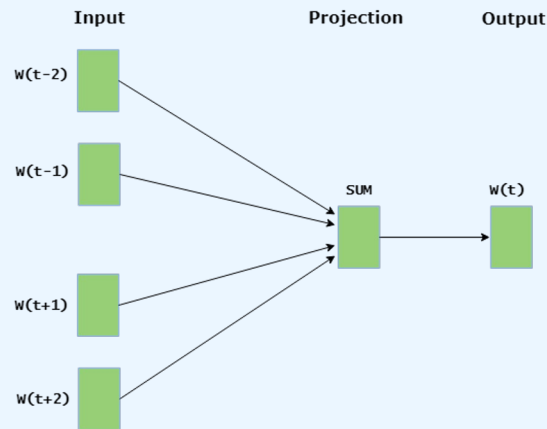
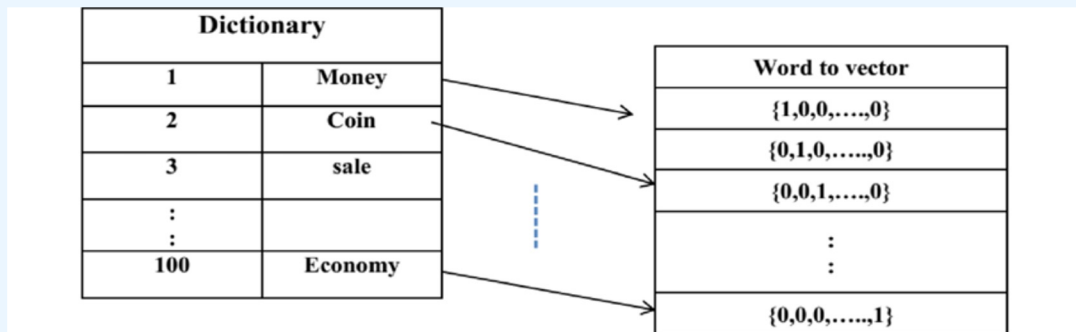
# Attention

- Decoder má informace o stavech Encoderu
- exponenciální složitost !

	I	love	you
Je	0.94	0.02	0.04
t'	0.11	0.01	0.88
aime	0.03	0.95	0.02

# Textová reprezentace

- Znak ASCII/UTF ?
  - Není vhodné
- One-hot encoding
  - Chybí význam slov
  - Problém skloňování/časování
  - Ontologie - nefunguje
- Word2Vec
  - Tomáš Mikolov
  - Vektorová reprezentace
  - Bezkontextové
  - Glove
- BERT, GPT-3
  - kontext
- BPE (Byte pair encoding)
  - tokeny



# Tokenizace

Tokens	Characters
224	504

Znaven tím vším, já chci jen smrt a klid,  
jen nevidět, jak žebrá poctivec,  
jak pýchou dme se pouhý parazit,  
jak pokřiví se každá čistá věc,  
jak trapně září pozlátko všech poct,  
jak dívčí cudnost brutálně rve chtič,  
jak sprostota se sápe na slušnost,  
jak blbost na schopné si bere bič,  
jak umění je pořád služkou mocných,  
jak hloupost zpupně chytrým poroučí,  
jak prostá pravda je všem prostě pro smích,  
jak zlo se dobru chechtá do očí.  
Znaven tím vším, já umřel bych tak rád,  
jen nemuset tu tebe zanechat.

Text

Token IDs

# Tokenizace

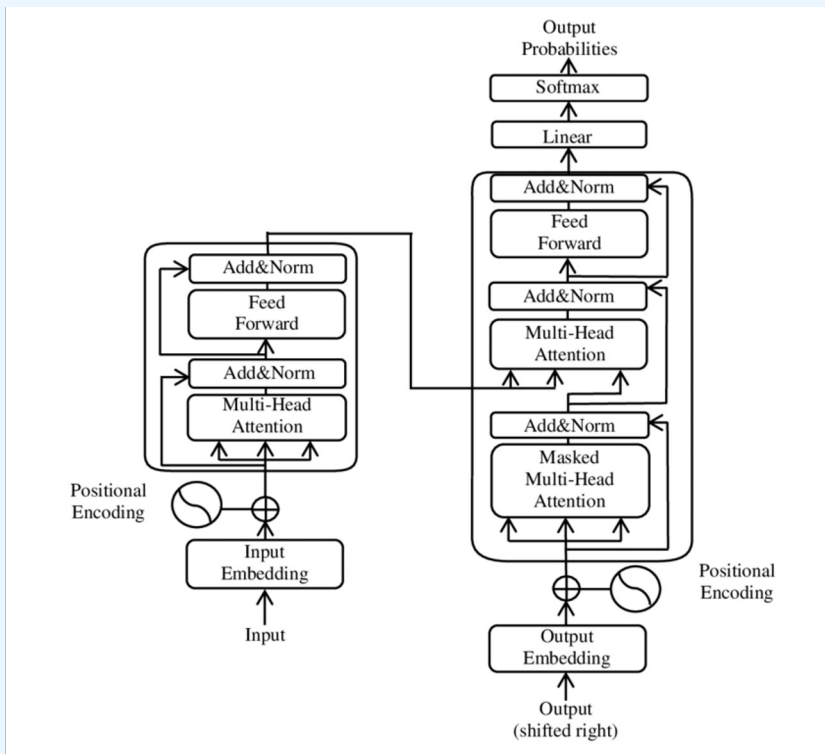
Tokens	Characters
134	568

Tired with all these, for restful death I cry,  
As, to behold desert a beggar born,  
And needy nothing trimm'd in jollity,  
And purest faith unhappily forsworn,  
And gilded honour shamefully misplaced,  
And maiden virtue rudely strumpeted,  
And right perfection wrongfully disgraced,  
And strength by limping sway disabled,  
And art made tongue-tied by authority,  
And folly doctor-like controlling skill,  
And simple truth miscall'd simplicity,  
And captive good attending captain ill:  
Tired with all these, from these would I be gone,  
Save that, to die, I leave my love alone.

Text  Token IDs

# Transformer

- GPT
- Problém paralelizace seq2seq
  - Zahodíme a necháme jen Attention



# Jazykový model

- Aplikace předpovídající další slovo v textu !!!
- Jednoměrné vs. obousměrné
- Použití
  - Speech to text
  - Prediktivní klávesnice
  - Pravopis
  - Překlady
  - ...
  - Nečekané vlastnosti !!!



# Limitace

- Velikost kontextového okna
  - Opakovat v promptu důležité informace
- Jazyk
  - Trénování
  - Tokenizace
- Systémový/uživatelský prompt
  - Prompt injection
- Je to stroj na hledání dalšího slova
  - Moduly
    - Python
    - Dall-E
    - Browser



# Limitace

- Jednoměrné vs. obousměrné
- Použití
  - Speech to text
  - Prediktivní klávesnice
  - Pravopis
  - Překlady
  - ...
  - Nečekané vlastnosti !!!

## ChatGPT

Abychom mohli spočítat, kolik kilometrů naběhá pes, můžeme to rozdělit do dvou částí:

1. Doba, po kterou pes běží, dokud myslivec nedojde do hájenky.
2. Celková vzdálenost, kterou pes naběhá za tuto dobu.

### Výpočet doby:

Myslivec jde z hospody do hájenky, což je 10 km, rychlostí 5 km/h. Doba, kterou myslivec stráví cestou, je tedy:

$$\text{čas} = \frac{\text{vzdálenost}}{\text{rychlost}} = \frac{10 \text{ km}}{5 \text{ km/h}} = 2 \text{ hodiny}$$

### Výpočet vzdálenosti, kterou naběhá pes:

Pokud pes běhá rychlostí 20 km/h a celou dobu, co myslivec jde domů (2 hodiny), pes také běhá, pak celková vzdálenost, kterou pes naběhá, je:

$$\text{vzdálenost} = \text{rychlost} \times \text{čas} = 20 \text{ km/h} \times 2 \text{ hodiny} = 40 \text{ km}$$

Takže pes během cesty myslivce z hospody do hájenky naběhá celkem 40 kilometrů.

## Úloha s psem a myslivcem

### Výpočet vzdálenosti:

Pes běhá tam a zpět, takže celková vzdálenost, kterou naběhá, se skládá ze dvou částí:

- Vzdálenost z hospody do hájenky a zpět:  $2 \times 10 \text{ km} = 20 \text{ km}$
- Vzdálenost, kterou pes uběhne v ústrety myslivci a zpět za celou cestu:

Myslivec jde rychlostí 5 km/h a pes běží 20 km/h, takže pes uběhne vzdálenost, kterou myslivec ušel, 4x rychleji.

To znamená, že pes uběhne  $4 \times 10 \text{ km} = 40 \text{ km}$  za celou cestu.

### Celková vzdálenost:

Celková vzdálenost, kterou pes naběhá, je součtem obou částí:  $20 \text{ km} + 40 \text{ km} = 60 \text{ km}$

### Odpověď:

Pes naběhne **60 kilometrů**.

### Poznámka:

V tomto řešení se předpokládá, že pes běží nepřetržitě stejnou rychlostí 20 km/h a že se nezastavuje, když potkává myslivce.

Ve skutečnosti se pes pravděpodobně bude zastavovat, hrát si a čenichat po cestě, takže skutečná vzdálenost, kterou naběhá, může být o něco menší.

## Kolik naběhá pes?

Nejprve si vypočítáme, kolikrát pes doběhne z hájenky do hospody a zpátky, než se myslivec dostane domů:

- Vzdálenost z hospody do hájenky je 10 km.
- Myslivec jde rychlostí 5 km/h, takže cesta mu trvá  $10 \text{ km} / 5 \text{ km/h} = 2$  hodiny.
- Za 2 hodiny pes doběhne z hájenky do hospody a zpátky  $2 * 2 = 4$  krát.

Nyní můžeme vypočítat celkovou vzdálenost, kterou pes uběhne:

- Jedna cesta z hájenky do hospody a zpátky je  $10 \text{ km} + 10 \text{ km} = 20 \text{ km}$ .
- Pes uběhne  $4 * 20 \text{ km} = 80 \text{ km}$ .

**Odpověď:** Pes naběhne **80 km**.

**Poznámka:** V tomto výpočtu jsme předpokládali, že pes běží stejnou rychlostí tam i zpátky a že se nezastavuje. Ve skutečnosti pes pravděpodobně poběží pomaleji, když se unaví, a může se také zastavovat na čůrání a pití. To by mohlo vést k tomu, že pes uběhne o něco méně než 80 km.

## Kolik pes uběhne?

**Nejprve spočítáme celkovou vzdálenost, kterou pes uběhne tam a zpátky:**

- Pes běží z hospody do hájenky a zpět za myslivcem, dokud se nesetkají. To znamená, že pes uběhne vzdálenost  $2 * 10 \text{ km} = 20 \text{ km}$ .
- Navíc pes běží naproti myslivcovi, když už je myslivec na cestě z hospody do hájenky. V tomto případě pes uběhne vzdálenost, kterou myslivec ušel, když se pes s ním poprvé setkal.

**Abychom zjistili, jak daleko pes uběhl, když se s myslivcem poprvé setkal, spočítáme, kolik kilometrů ušel myslivec v tu dobu:**

- Myslivec jde rychlostí 5 km/h a vzdálenost z hospody do hájenky je 10 km. To znamená, že myslivec ušel  $5 \text{ km/h} * 1 \text{ h} = 5 \text{ km}$ .

**Nyní můžeme spočítat celkovou vzdálenost, kterou pes uběhl:**

- Pes uběhl  $20 \text{ km}$  tam a zpátky +  $5 \text{ km}$ , když se s myslivcem poprvé setkal =  $25 \text{ km}$ .

**Odpověď:** Pes uběhne celkem 25 km.

**Poznámka:** V tomto výpočtu předpokládáme, že pes běží stejnou rychlostí tam i zpátky a že se s myslivcem setká přesně v polovině cesty. V reálném světě by se pes mohl pohybovat různou rychlostí a setkání s myslivcem by se mohlo stát v jiném bodě cesty, čímž by se celková vzdálenost, kterou pes uběhne, mírně lišila.

# DĚKUJI!

Nějaké dotazy ?

benes@netvia.cz

+420 724 582 980

www.netvia.cz



**CREDITS:** This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)

