

Využití jazykových modelů pro získávání informací

Pavel Kraus

540455@mail.muni.cz

Ústav výpočetní techniky, Masarykova univerzita

20. listopadu 2024

Osnova prezentace

1. Obecné představení jazykových modelů
2. Využití v různých oblastech
3. Ukázka získávání informací z PDF dokumentu
4. Limity využívání jazykových modelů

Co jsou to jazykové modely

- Jazykový model = **nástroj**
- K čemu takový nástroj můžeme využít? Generování textu, překlady textů... a pro **získávání informací**
- Jak se s jazykovými modely pracuje?
 - Stačí znát lidský jazyk
 - Použít populární platformy jako chatbot ChatGPT (chatgpt.com)
 - Chatbotovi zadáváme instrukce, které chatbot plní
- Jazykový model není jen jeden jediný

Získávání informací

Obecně známé informace

- *"Vysvětli mi, proč je CO2 problém pro lidstvo."*
- *"Jak vznikají hlubokomořské příkopy?"*

Získávání informací z dat, které jazykovému modelu poskytnu.

- Jakákoliv nestrukturovaná data
 - Textové soubory - PDF, Word...
 - Obrázky
 - Zvukové soubory
 - Webové stránky
- Soukromá data

Příklady získávání informací z dat

Příklad 1: PDF dokument s informacemi o Green Deal

- *"Vypiš mi, s jakými čísly tenhle dokument pracuje"*
- *"Co tento dokument říká o obnovitelných zdrojích?"*

Příklad 2: Databáze všech publikací na Masarykově univerzitě

- *"Shrň klíčové body všech publikací na téma umělé inteligence."*
- *"Které publikace se zabývají změnou klimatu?"*

Příklad 3: Databáze mapových podkladů

- *"Na kterých mapách je Bratislava?"*

Ukázka získávání informací z PDF dokumentu

- PDF dokument GREEN DEAL: Fakta - čísla - dopady
- Volně dostupný nástroj ChatGPT
- Budeme chtít vědět:
 - *"Vypiš mi, s jakými čísly tenhle dokument pracuje"*
 - *"Co tento dokument říká o obnovitelných zdrojích?"*


S čím ti můžu pomoci?

Napsat zprávu pro ChatGPT



Připojit soubory




 Vytvořit obrázek

 Porad' mi

 Kód

 Vytvoř plán

 Analyzovat obrázky

Více

S čím ti můžu pomoci?



GREEN DEAL Fakta - čísla - dopady.pdf
PDF



Vypiš mi, s jakými čísly tenhle dokument pracuje



 Vytvořit obrázek

 Porad' mi

 Kód

 Vytvoř plán

 Analyzovat obrázky

Více



GREEN DEAL Fakta - čísla - dopady.pdf
PDF

Vypiš mi, s jakými čísly tenhle dokument pracuje



Dokument pracuje s následujícími čísly a fakty týkajícími se Green Dealu:

1. **260 miliard eur ročně:** Dodatečné investice potřebné k dosažení cílů Green Dealu do roku 2030 (1,5 % HDP roku 2018).
2. **Snížení emisí skleníkových plynů do roku 2030:** O nejméně 50 %, ideálně 55 %, oproti roku 1990.
3. **30 % chráněné půdy a moří EU:** Cíl ochrany biologické rozmanitosti do roku 2030.
4. **Snížení pesticidů:** O 50 % do roku 2030.
5. **Zastavení úbytku včel a výsadba stromů:** 3 miliardy stromů do roku 2030.
6. **40 % spotřeby energie a 36 % emisí v EU:** Pochází z budov, přičemž pouze 1 % budov projde ročně renovací.

Co tento dokument říká o obnovitelných zdrojích?



Dokument se věnuje obnovitelným zdrojům energie v několika částech a zmiňuje:

Hlavní cíle:

1. Podpora energie z obnovitelných zdrojů:

- Snížení emisí skleníkových plynů prostřednictvím rozšíření obnovitelných zdrojů energie.
- Integrace energie z obnovitelných zdrojů do budov a jejich vytápění/chlazení.

2. Přezkum směrnice o obnovitelných zdrojích energie:

- Do června 2021 byl plánován přezkum, který měl zvážit zpřísnění stávajících cílů, zejména v oblasti vytápění a chlazení.

3. Zvýšení podílu obnovitelných zdrojů:

- EU plánuje postupnou dekarbonizaci energetiky, která má hrát klíčovou roli při dosažení klimatické neutrality do roku 2050.

Limity využívání jazykových modelů

- Získávání informací z velkých souborů dat není vždy jednoduché
- Modely (zatím) nepracují nejlépe v některých specifických doménách (například zdravotnictví)
- Modely mohou generovat nesmyslné odpovědi
- Používání veřejných nástrojů znamená sdílení dat s cizí společností

Děkuji za vaši pozornost!

**MASARYKOVA
UNIVERZITA**