

ChatGPT in umetna inteligenca v izobraževanju

Gimnazija Jožeta Plečnika Ljubljana

Ljupčo Todorovski

Univerza v Ljubljani, Fakulteta za matematiko in fiziko
Institut Jožef Stefan, Odsek za tehnologije znanja (E8)

Junij 2023

- 1 Kaj je ChatGPT?
- 2 Uporaba ChatGPT v izobraževanju

Jezikovni modeli

So nevronske mreže

- Po strukturi enostavne, a ogromne
- Sestavljene iz milijard povezanih enostavnih enot, nevronov
- Na osnovi poizvedb uporabnika lahko tvorijo besedila
- Tudi slike in videe

Primer uporabe jezikovnih modelov: storitev Googlov prevajalnik

Nevronske mreže

Sestavljene so iz nevronov in sinaps

- **Nevron** izračuna svoje stanje iz stanja z njim povezanih nevronov
- In uteži na **sinapsah, ki nevrone medsebojno povezujejo**
- Stanje vhodnih nevronov so podatki, npr. vprašanje od uporabnika
- Stanja izhodnih nevronov postrežejo uporabniku podatke, odgovor

Učenje nevronske mreže

Uteži na sinapsah so na začetku (pred učenjem) neznane

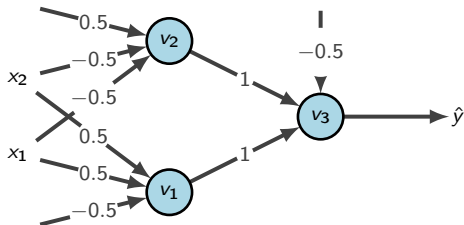
- Med učenjem, nevronske mreže kažejo **veliko učnih primerov**
- Računalnik med učenjem **nastavi uteži na sinapsah** tako
- Da nevronska mreža čim boljše **poustvari učne primere**, izhode

Perceptron ima omejeno kapaciteto: Ekskluzivni ali

Podatki za učenje modela

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

Za modeliranje rabimo *večplastno* nevronska mrežo (Werbos 1975)



ChatGPT

Je klepetalni agent, ki uporablja jezikovni model GPT

- Slednji je, kot smo videli prej, ogromna nevronska mreža
- Naučena iz **vseh besedil** dostopnih podjetju OpenAI
- Zato postopek učenja terja **ogromno časa in računalniške moči**
- Nevronska mreža modela GPT-3.5 ima **več sto milijard sinaps**
- Naučena je iz besedil, ki vsebujejo **več sto milijard besed**

- Različica GPT-3.5 v prosto dostopnem ChatGPT
- Različica GPT-4 v plačljivi verziji ChatGPT+

Je ChatGPT umetna inteligenca?

Kaj sploh je umetna inteligenca (UI)?

Alan Turing, članek iz leta 1950

- Naslov in centralno vprašanje: Ali stroji (računalniki) lahko mislijo?
- **Igra imitacije**, Turingov test, ki je dolga leta opredeljeval UI

Torej, je ChatGPT umetna inteligenca?

Ja, ChatGPT je **šibka** umetna inteligenca.

Tri ravni umetne inteligence

- Šibka UI** zmožna reševanja zelo konkretnih nalog, ki jih poda človek
- Močna UI** zmožna razumevanja, učenja in splošne uporabe znanja za reševanje poljubnih nalog (nevronske mreže temu niso kos)
- Super UI** superiorna vsem vidikom človeške inteligence

Je UI koristna?

Nedvomno je koristna

- Šibka UI nam lahko pomaga hitreje reševati probleme
- Omogoča **hiter in enostaven dostop** do ogromne količine znanja
- Številne možnosti za **pospešek** znanstvenega raziskovanja in poslovanja

Nevarnosti in težave: mitske in resnične

Mit: UI predstavlja eksistencialno tveganje za človeštvo

- Šibka UI ni avtonomna, ne more se sama odločati
- Povzroča nepotrebno paniko med splošno populacijo
- Odvrča nas od smotrne uporabe UI in jezikovnih modelov
- Kljub hitremu razvoju, smo še zelo daleč od močne ali super UI

Mit je lahko nevaren, saj odvrča pozornost od resničnih nevarnosti UI.

Resnično pomembne težave in vprašanja: Vpliv na družbo

Vpletanje UI v demokratične procese

- Spomnimo se vloge socialnih omrežij v kampanji Brexit
- Takrat so manipulacije javnega mnenja izvajali najeti ljudje
- ChatGPT lahko zelo učinkovito manipulira in ustvarja lažne novice

Širjenje predsodkov in diskriminacije

- Povzame predsodke in diskriminacije v besedilih za učenje UI
- In jih zelo učinkovito ter na široko poustvarja in razširja

Resnično pomembne težave in vprašanja: Nepreglednost

Slabo razumevanje in nepreglednost delovanja ChatGPT

- Zaradi ogromne velikosti nevronske mreže
- **Ne znamo dobro pojasniti in razumeti njihovo delovanje**
- Zato ga tudi težje nadzorujemo

Nagnjenost k napakam in izmišljanju

- ChatGPT **ne razume besedil** in zato nima odnosa do njihove resničnosti
- V jezikovni model integrira tudi **neresnice in zavajanja v učnih podatkih**

Resnično pomembne težave in vprašanja: Širši vplivi

Zloraba intelektualne lastnine

- Jezikovni modeli se učijo iz **avtorskih** besedil
- Pripadajo človeštvu, ne zgolj podjetjem, ki razvijajo jezikovne modele
- Kako bodo ta podjetja kompenzirala avtorske pravice?

Vpliv na delovno okolje in ustvarjanje

- **Izkoriščanje delavcev** za pripravo učnih podatkov za UI
- **Nadomeščanje delovnih mest** in celih poklicev

Delovanje jezikovnih modelov terja **velikanski ogljični odtis**.

Edini racionalni odgovor: Regulacija razvoja in uporabe

Regulacija s politikami

- Definirajo pojem **zaupanja vredne umetne inteligence**
- Zahteve po razvoju UI, ki **lahko razloži svoje odločitve**
- Debata o omejevanju uporabe jezikovnih modelov in ChatGPT

- Tehnološki razvoj močno prehiteva javno debato o regulaciji

Regulacija na področju izobraževanja

Tukaj lahko ukrepamo pedagogi sami, debata v nadaljevanju.

Koncept odgovorne uporabe orodij UI

Zaenkrat jo večinoma **definirajo uredništva znanstvenih revij**.

Principa odgovorne uporabe

- 1 Prezemanje **popolne odgovornosti** za napisano besedilo
- 2 Iskrenost in **transparentnost pri navajanju virov**

Princip popolne odgovornosti

ChatGPT ne sme biti naveden med avtorji besedila

- Avtorji smo lahko zgolj ljudje
- S tem ne prenašamo odgovornosti na UI in ChatGPT
- Jamčimo, da **smo preverili pravilnost samodejno ustvarjene vsebine**
- Preprečimo ustvarjanje neresnic za nadaljnje učenje jezikovnih modelov

Princip transparentnosti uporabe

V posebnem dodatku dela razkrijemo interakcijo s ChatGPT

- **Obvezno za bistvene dele članka** z vsebinskimi prispevki
- Manj bistveno za uvodne dele članka in npr. pregled obstoječega dela

Programska orodja za preverjanje vsebine

- **Orodja za kontrolo** upoštevanja principa **transparentnosti**
- Identificirajo dele besedila, ki so bili samodejno ustvarjeni z UI
- Pozor: Težave podobne tistimi z orodji protivirusne zaščite

Praktični napotki: Uporaba, začetek

NEMUDOMA se navadite uporabljati ChatGPT

- Brezplačna prijava na naslovu <https://chat.openai.com/>
- Odpreti morate novi uporabniški račun (zamudno, ne priporočam)
- Ali uporabite obstoječi račun, npr. Googleov (takoj, priporočam)

- Različica ChatGPT+, ki sloni na modelu GPT-4 je plačljiva
- Zaenkrat ni na voljo kolektivnih naročnin, zgolj za posameznike
- Trenutni strošek mesečne naročnine 20 USD

Praktični napotki: Preverjanje in kontrola

Začnite uporabljati orodje GLTR

- Brez registracije na naslovu <http://gltr.io/>
- Pobarva besedilo tako, da zelena barva nakazuje uporabo UI
- Rdeče ali vijolično pobarvano besedilo nakazuje manj UI

Pozor: Slabosti

- Meri **stopnjo pričakovanosti** besednih zvez
- Orodje je zastarelo, razvili so ga leta 2019 v času GPT-2
- Zaradi tega velikokrat ne prepozna samodejno (z UI oz. ChatGPT) ustvarjenega slovenskega besedila

Preizkusite tudi <https://platform.openai.com/ai-text-classifier>.

Primer uporabe ChatGPT: Vprašanja za test

Iztočnica (besedilo, ki ga vnesete Vi)

Sestavi mi prosim odstavek z manjkajočimi besedami za test o jezikovnih modelih in ChatGPT.

Odgovor (besedilo, ki ga tvori ChatGPT)

Jezikovni modeli so izjemno uporabno orodje v ___ raziskavah in praksi. En tak model, imenovan ChatGPT, je razvit na arhitekturi ___, ki omogoča visoko kakovostno obdelavo naravnega jezika. ChatGPT je treniran na obsežnih korpusih ___, ki mu omogočajo izjemno raznolikost in obsežnost jezikovnega znanja. Model ChatGPT ima sposobnost ___ s uporabniki, odgovarjati na njihova vprašanja in nuditi pomoč pri iskanju informacij. Poleg tega ChatGPT lahko generira ___ besedilne vsebine, ki so ustrezne in kreativne. Kljub temu, da ChatGPT izkazuje impresivne jezikovne sposobnosti, mu manjka ___ in lastno razumevanje, saj temelji na vnaprej določenih modelih in podatkih.

Primer uporabe ChatGPT: Vprašanja za test

Iztočnica

Sestavi mi prosim vprašanje z izbiro pravega odgovora o uporabi drugega sklona v slovenščini.

Odgovor

Katera izmed naslednjih trditev pravilno prikazuje uporabo drugega sklona v slovenščini?

- a "Vidim mojega prijatelja v parku."
- b "Kupil sem knjigo v trgovini."
- c "Napišem pismo svojemu bratu."
- d "Zanima me tvojega novega avtomobila."

Pravilen odgovor: (c) "Napišem pismo svojemu bratu."

Primer uporabe ChatGPT: Matematična naloga (1)

Iztočnica

Sestavi mi prosim besedilno matematično nalogo, ki zahteva reševanje kvadratne enačbe.

Odgovor

Skupina prijateljev se odloči obiskati zabavišni park. Cena vstopnice za vsakega odraslega je 25 evrov, medtem ko je cena otroške vstopnice 15 evrov. Skupaj so prijatelji kupili 20 vstopnic in plačali 410 evrov. Prijatelji se sprašujejo, koliko odraslih in koliko otrok je v skupini.

Vprašanje: Sestavi kvadratno enačbo, ki opisuje situacijo, ter ugotovi, koliko odraslih in koliko otrok je v skupini.

Primer uporabe ChatGPT: Matematična naloga (2)

Iztočnica

Oprosti, ampak ta naloga zahteva reševanje sistema linearnih enačb, ne pa kvadratne.

Odgovor

Oprostite za napako. Tukaj je primerna naloga, ki zahteva reševanje kvadratne enačbe: Vrtnar Janez je opazil, da mu nekaj zajcev poškoduje rastline v njegovem vrtu. Odločil se je, da bo postavil ograjo okoli vrta, da bi preprečil vstop zajcem. Želi postaviti kvadratno ograjo s stranico dolžine 12 metrov. Vendar nima dovolj ograjnega materiala in bo moral uporabiti samo 36 metrov ograjnega materiala. Janez se sprašuje, kakšno naj bo največje območje ograjenega vrta, ki ga lahko postavi s tem omejenim materialom. Vprašanje: Sestavite in rešite kvadratno enačbo, da ugotovite največje območje ograjenega vrta, ki ga Janez lahko postavi s 36 metri ograjnega materiala.

Primer uporabe ChatGPT: Matematična naloga (3)

Iztočnica

Tudi tukaj ne rabimo formule za reševanje kvadratne enačbe, ta naloga še vedno ni dobra.

Odgovor

Družina se odloči postaviti bazen v obliki pravokotnika v svojem dvorišču. Želijo, da je dolžina bazena za 3 metre daljša od širine. Skupna površina bazena naj bo 84 kvadratnih metrov. Družina se sprašuje, kakšni naj bosta dolžina in širina bazena.

Vprašanje: Sestavi kvadratno enačbo, ki opisuje situacijo, in reši jo, da ugotoviš dolžino in širino bazena.

Primer uporabe ChatGPT: Priprava predavanja

Iztočnica

Sestavi mi prosim 10 prosojnic za predavanje “Uvod v ChatGPT”.

Izseki (daljšega) odgovora

Prosojnica 1: Naslov: Uvod v ChatGPT

- Predstavitvev ChatGPT in njegovega pomena v naravnojezikovnem procesiranju
- Zgodovinski pregled razvoja jezikovnih modelov

Prosojnica 2: Naslov: Kako deluje ChatGPT?

- Razlaga arhitekture jezikovnega modela GPT
- Vlogo globokih nevronskih mrež pri generiranju besedila

Prosojnica 6: Naslov: Izzivi pri uporabi ChatGPT

Prosojnica 7: Naslov: Etika in odgovornost pri uporabi ChatGPT

Prosojnica 10: Naslov: Nasveti za učinkovito uporabo ChatGPT

Primerna in neprimerna uporaba

Primerna uporaba

- Kot generator idej za vprašanja in naloge
- A tudi za predavanja ali priprave (poskusite sami!)
- Za preverjanje težavnosti domačih nalog ali testov
- Preverite ali ChatGPT pozna odgovore na vaša vprašanja

Neprimerna uporaba

Za preverjanje točnosti izjav ali preverjanje vira besedila (v nadaljevanju).

ChatGPT se pogosto laže

Iztočnica

Ali je besedilo "Družina se odloči postaviti bazen v obliki pravokotnika v svojem dvorišču. Želijo, da je dolžina bazena za 3 metre daljša od širine. Skupna površina bazena naj bo 84 kvadratnih metrov. Družina se sprašuje, kakšni naj bosta dolžina in širina bazena." napisal človek ali ChatGPT?

Odgovor

Besedilo, ki ga omenjate, je bilo napisano s strani človeka in ne zgenerirano s strani ChatGPT.

Pozor: gre za besedilo naloge, ki jo je sestavil prej (prosojnica 25).

In se tudi prostodušno skesa

Iztočnica

Na osnovi česa to trdiš?

Odgovor

Oprostite za zavajajoč odgovor v prejšnjem sporočilu. Kot jezikovni model GPT-3.5, nimam vpogleda v to, kdo je napisal določeno besedilo. Glede na značilnosti besedila, kot je struktura in jezikovni slog, je **težko z gotovostjo trditi, ali je bilo besedilo napisano s strani človeka ali generirano s pomočjo jezikovnega modela**. Za določitev avtorstva bi bilo potrebno imeti dodatne informacije o izvoru besedila.

Namesto zaključka

- Začnite uporabljati ChatGPT, <https://chat.openai.com/>
- Seznanjajte se z orodji za preverjanje vira, <http://gltr.io/>
- Preizkusite tudi <https://platform.openai.com/ai-text-classifier>

- Vadimo pisanje iztočnic, preizkušajmo svoje ideje za naloge in teste
- Učimo se skupaj odgovorne uporabe jezikovnih modelov in UI
- Poučujmo naše dijake in študente o odgovorni, kritični uporabi UI

- Iščimo pravo razmerje med spodbujanjem uporabe UI
- In prepovedmi v tistih delih učnega procesa, ker te smotrne