

Izdelava in demonstracija modelov za zdravorazumsko sklepanje za slovenščino

Erik Novak

Odsek za umetno inteligenco

Institut “Jožef Stefan”

Pregled predstavitve

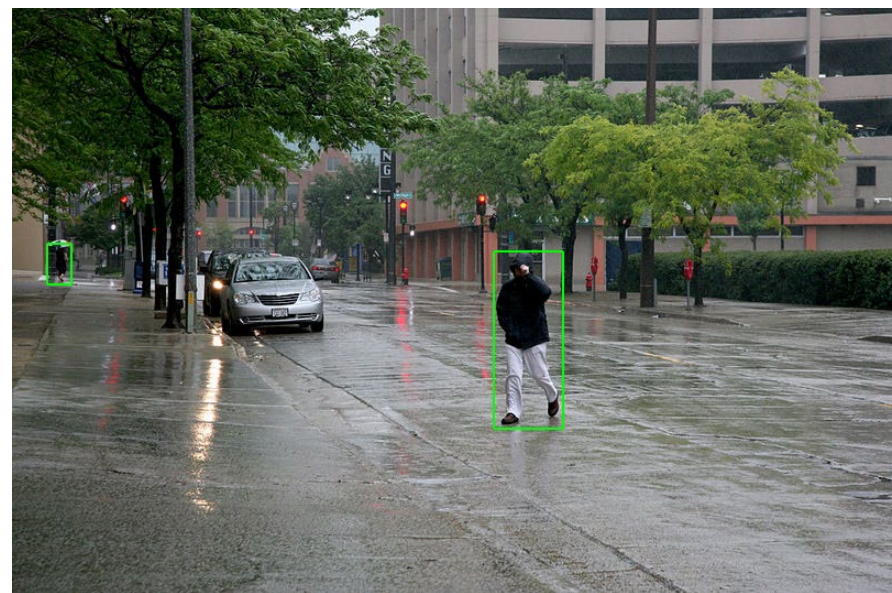
- Kaj je zdravorazumsko sklepanje?
- Zgodovina razvoja zdravorazumskega sklepanja v UI
- Jezikovni modeli
- Podatki
- Razvoj modela za zdravorazumsko sklepanje za slovenščino
- Rezultati
- Praktična uporaba
- Naslednji koraki
- Zaključek

Kaj je zdravorazumsko sklepanje?

Zdravorazumsko sklepanje se nanaša na sposobnost napovedovanja verjetnih zaključkov, ki izhajajo iz izbranega znanja.

Primeri raziskav na področju zdravorazumskega sklepanja v umetni inteligenci

- **Slike in posnetki.** Razumevanje postavitve predmetov v sliki oz. posnetku
- **Besedilo.** Razumevanje in sklepanje iz besedila



Zgodovina razvoja zdravorazumskega sklepanja v UI

- Priprava taxonomij in enciklopedije znanja

- Fokus na znanju o “kaj”, ne pa znanju o “kako” in “zakaj”

Ernest Davis and Gary Marcus. 2015. Commonsense reasoning and commonsense knowledge in artificial intelligence. Commun. ACM 58, 9 (September 2015), 92–103.

- Izluščanje zdravorazumskih sklepnih pravil iz različnih virov podatkov

Luis Antonio Galárraga, Christina Teflioudi, Katja Hose, and Fabian Suchanek. 2013. AMIE: association rule mining under incomplete evidence in ontological knowledge bases. In Proceedings of the 22nd international conference on World Wide Web (WWW '13). Association for Computing Machinery, New York, NY, USA, 413–422.

Yang, Bishan & Yih, Wen-tau & He, Xiaodong & Gao, Jianfeng & Deng, li. 2014. Embedding Entities and Relations for Learning and Inference in Knowledge Bases.

- Uporaba generativnih jezikovnih modelov

A. Bosselut, H. Rashkin, M. Sap, C. Malaviya, A. Celikyilmaz, and Y. Chai, 2019. COMET: Commonsense Transformers for Automatic Knowledge Graph Construction. in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 4762–4779.

J. D. Hwang, C. Bhagavatula, R. Le Bras, J. Da, K. Sakaguchi, and A. Bosselut, 2021. (Comet-) Atomic 2020: On Symbolic and Neural Commonsense Knowledge Graphs. Proceedings of the AAAI Conference on Artificial Intelligence, 6384–6392.

Jezikovni modeli

Cilj jezikovnih modelov

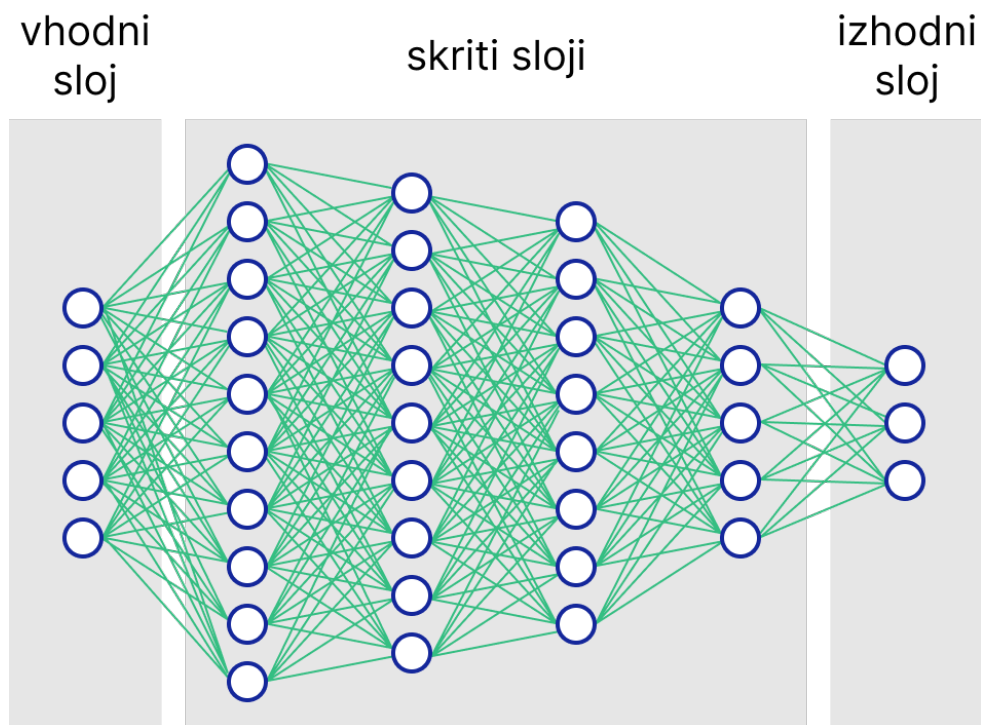
Jezikovni modeli

- Jezikovni model napove s kolikšno verjetnostjo se beseda pojavi v danem kontekstu
 - Primer: “Danes je resnično _____ dan.”
- V preteklosti so jezikovni modeli delovali na osnovi jezikovnih in besednih statistikah
- Danes delujejo na osnovi rekurzivnih nevronske mreže in **transformer arhitekture**



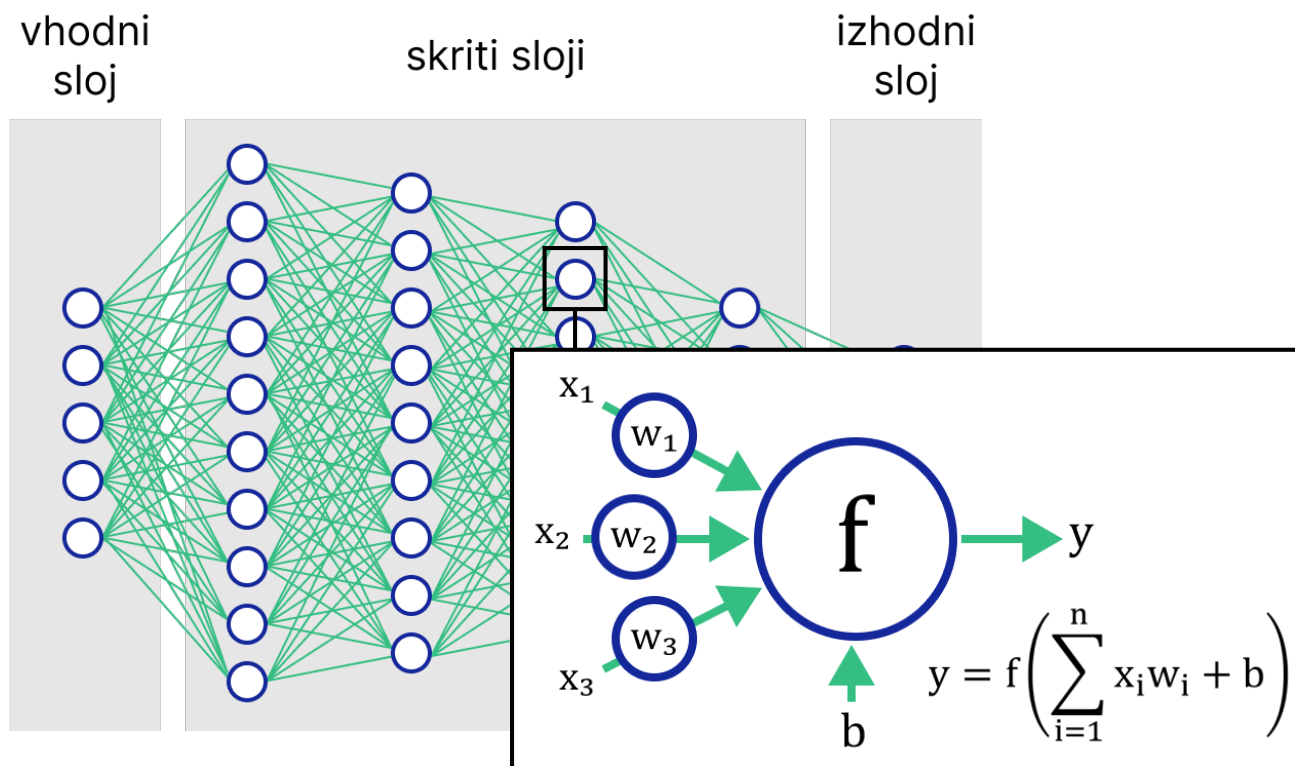
Neuronske mreže

Jezikovni modeli



Neuronske mreže (nadaljevanje)

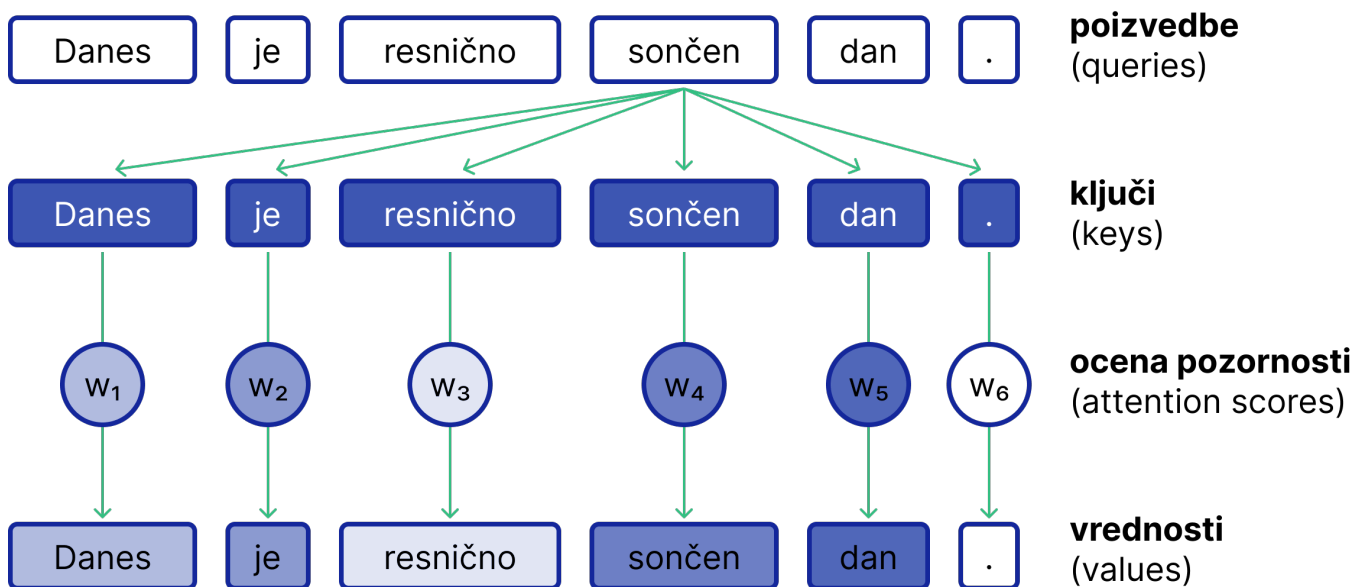
Jezikovni modeli



Mehanizem pozornosti

Jezikovni modeli

Mehanizem pozornosti omogoča pregled nad celotnim stavkom in selektivno izbiranje informacij, ki so potrebne v trenutnem koraku napovedovanja.

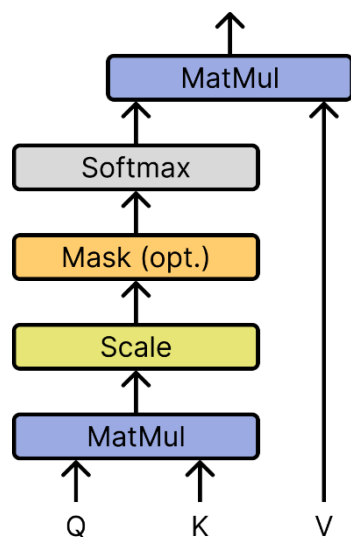


Več-glava pozornost

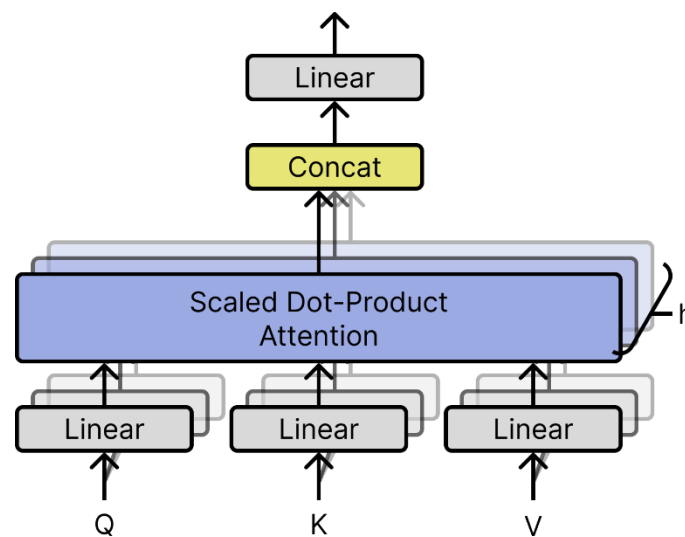
Jezikovni modeli

Več-glava pozornost izračuna več različnih ocen pozornosti (angl. attention scores), kar omogoča izračun različnih reprezentacij besed.

- Q poizvedbe
- K ključi
- V vrednosti



mehanizem pozornosti



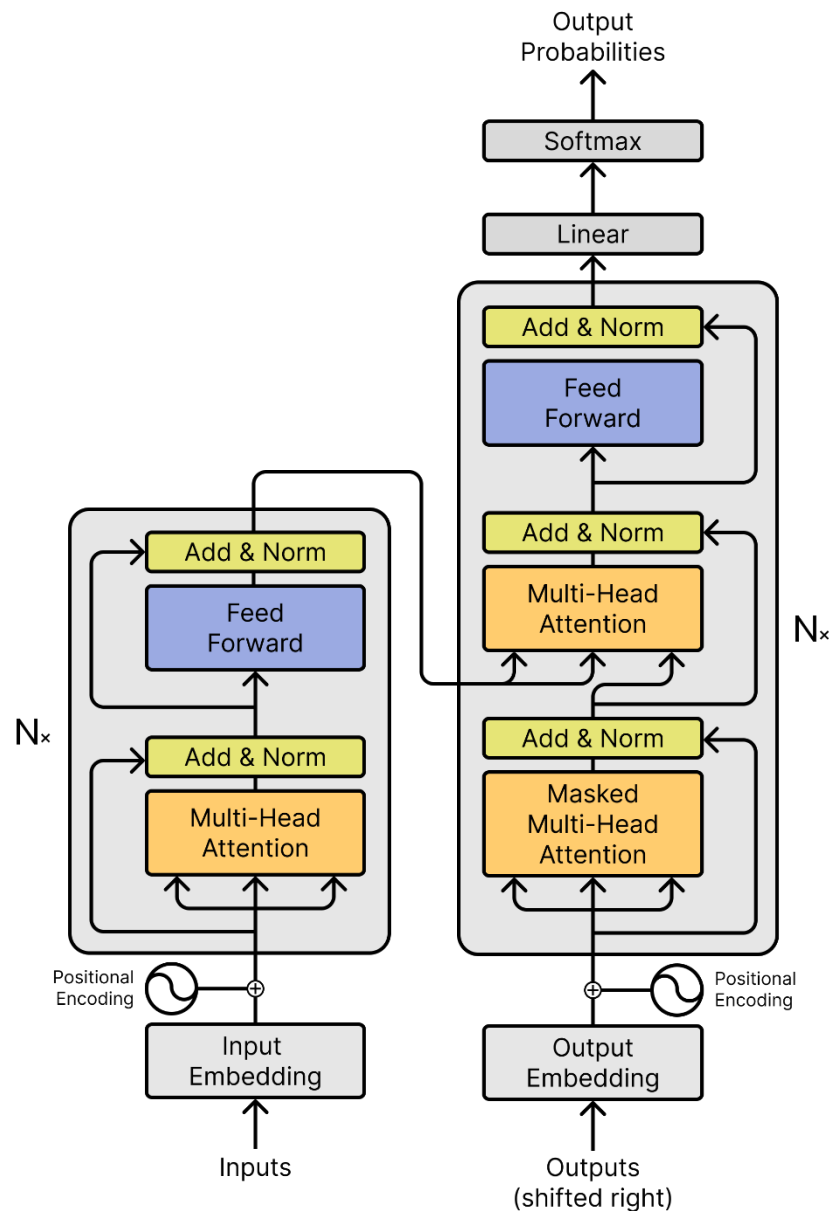
več-glava pozornost

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

Transformer model

Jezikovni modeli

- Transformer modeli sledijo encoder-decoder arhitekturi
- Vhodni podatki so vektorske predstavitev besed vhodnega teksta
- Del vhodnih podatkov v decoder so skriti - namen je da model ne vidi besed, ki jih še ni napovedal, s katerimi bi lahko “goljufal”

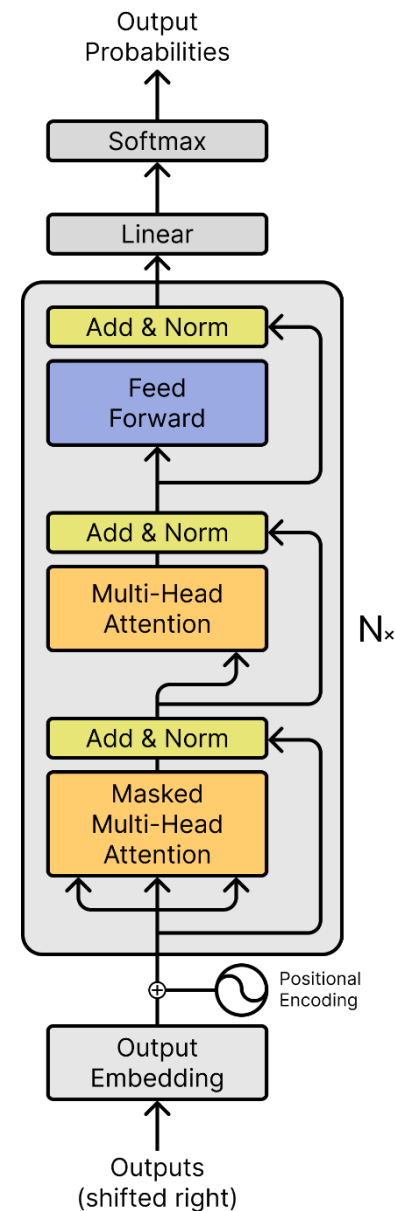


Družina GPT jezikovnih modelov

Jezikovni modeli

- Arhitektura: transformer decoder
- Delno nadzorovan pristop za reševanje nalog povezanih z razumevanjem jezikov
- Kombinira nenadzorovano pred-treniranje in nadzarovano fino nastavitvev

Model	Število parametrov
GPT	125M
GPT-2	1542M
GPT-3	175B
GPT-4	175B-280B (neuradno)

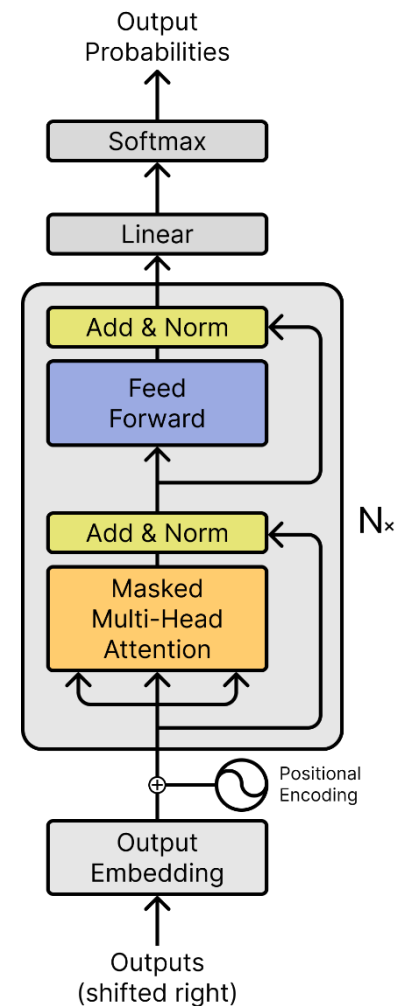


Družina GPT jezikovnih modelov

Jezikovni modeli

- Arhitektura: transformer decoder
- Delno nadzorovan pristop za reševanje nalog povezanih z razumevanjem jezikov
- Kombinira nenadzorovano pred-treniranje in nadzarovano fino nastavitvev

Model	Število parametrov
GPT	125M
GPT-2	1542M
GPT-3	175B
GPT-4	175B-280B (neuradno)



GPT-2 jezikovni model

Jezikovni model

- Naslednik modela GPT
 - Struktura modela podobna GPT z manjšimi popravki
- Generalni sistem za reševanje več različnih nalog brez spreminjanja oblike vhodnih podatkov
 - Pogoji niso več samo na vhodne podatke ampak tudi na nalogo ki jo mora model opraviti: $P(\text{output} \mid \text{input}, \text{task})$
- Primeri nalog in vhodnih podatkov:
 - **Prevajanje.** “Translate to French: English text, **French text**”
 - **Odgovarjanje na vprašanja.** “Answer the question: document text, question, **answer**”
 - **Zdravorazumsko sklepanje.** “Head sentence, relation, **reasoning**”



Slovenski GPT modeli

Jezikovni modeli

Tabela slovenskih GPT modelov ki smo jih uporabili pri raziskavi

Model	arhitektura	število parametrov	velikost vokabularja
<i>Modeli pripravljene na FRI</i>			
cjvt/gpt-sl-base	GPT-2	131,966,208	60,032
gpt-janez	GPT-J	134,275,387	32,059
gpt-janezek	GPT-J	31,742,779	32,059
<i>Ostali modeli</i>			
macedonizer/sl-gpt2	GPT-2	125,820,672	52,000
<i>Originalni ENG model</i>			
GPT2-XL	GPT-2	1,558,000,000	50,257

Natrenirani na slovenskih korpusih:

- Gigafida
- KAS
- slWaC
- MaCoCu

Podatki

Originalni podatki

Podatki

- Množica $ATOMIC_{20}^{20}$ je narejena z uporabo AWS Mechanical Turk
- Vsebuje 1,33M primerov
- Primeri porazdeljeni v 23 različnih relacij

Struktura primerov	HEAD	relacija	TAIL
	PersonX accidentally poured	xWant	to clean the mess

tip relacij	seznam vseh relacij
Povezani s fizičnimi predmeti in lastnostmi	ObjectUse, AtLocation, MadeUpOf, HasProperty, CapableOf, Desires, Not Desires
Centrirani na dogodke	IsAfter, HasSubEvent, IsBefore, HinderedBy, Causes, xReason, isFilledBy
Povezane s socialnimi interakcijami	xNeed, xAttr, xEffect, xReact, xWant, xIntent, oEffect, oReact, oWant

Podatki so v angleščini.
Potreben prevod!

Prevod podatkov

Podatki



DeepL

- Uporaba DeepL prevajalnika
 - Kvalitetno generiranje avtomatskih prevodov za slovenščino
 - Dostop in uporaba preko APIja
- Ločeno pošiljanje delov besedil za prevajanje
 - Relacij ne prevajamo

HEAD	relacija	TAIL
PersonX accidentally poured	xWant	to clean the mess

English ↕ Slovenian

PersonX accidentally poured ×

OsebaX je pomotoma izlila

Alternatives:

- OsebaX je pomotoma polila
- OsebaX pomotoma zlije
- OsebaX je pomotoma vlila



Ročno preverjanje prevodov

Podatki

- Avtomatski prevodi vsebujejo napake

tip besedila	original v angleščini	avtomatsko prevedeno besedilo	ročno popravljeno besedilo
HEAD	PersonX accidentally poured	OsebaX je pomotoma izlila	OsebaX je pomotoma polila
HEAD	PersonX abuses PersonX's power	OsebaX zlorablja pooblastila osebeX	OsebaX zlorablja moč osebeX
TAIL	to clean clothes	za čiščenje oblačil	očistiti oblačila

- Potrebno ročno preverjanje in popravljanje prevodov
 - Sistem za ročno preverjanje prevodov razvil [Odsek za računalništvo in informacijske tehnologije](#) na [FERI, Maribor](#)

Ročno preverjanje prevodov (nadaljevanje)

Podatki

Urejanje prevodov [Datoteka train.tsv](#) Možnosti

Seznam prevodov

[Prejšnja](#) Stran 4 od 10 [Naslednja](#) [Skoči na stran](#) Stran [Ei. na strani](#) Ei #

Vsi (neurejeni) prikazani so OK

Samo neurejeni Filterjaj

	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oEffect	TAIL: PersonX stutters with nervousness (izvirnik) Osebak od nervevo zaječe (prevod)	▼
	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oEffect	TAIL: receives kindness (izvirnik) prejema prijaznost (prevod)	▼
OK	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeX (prevod)	RELATION: oSent	TAIL: to be thanked. (izvirnik) za zahvalo. (prevod)	▼
OK	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oEffect	TAIL: stresses (izvirnik) obremenjuje (prevod)	▼
OK	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oEffect	TAIL: PersonY's stress reduces so there muscles relax (izvirnik) Stres se zmanjša, zato se mišice sprostitjo. (prevod)	▼
OK	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oEffect	TAIL: gets help (izvirnik) dobi pomoč (prevod)	▼
OK	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oEffect	TAIL: becomes relieved of responsibility (izvirnik) je razbremenjen odgovornosti. (prevod)	▼
OK	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oIntent	TAIL: to help a friend (izvirnik) pomagati prijatelju (prevod)	▼
OK	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oEffect	TAIL: commits a kind act (izvirnik) stori prijazno dejanje. (prevod)	▼
OK	HEAD: PersonX acts on PersonY's behalf (izvirnik) Osebak deluje v imenu osebeY (prevod)	RELATION: oEffect	TAIL: helps a friend (izvirnik) pomaga prijatelju. (prevod)	▼

Ročno preverjanje prevodov (nadaljevanje)

Podatki

Če je prevod v redu

OK HEAD: PersonX acts on PersonY's behalf (izvirnik)
osebaX deluje v imenu osebeY (prevod) RELATION: xIntent TAIL: to help (izvirnik)
pomagati (prevod)

Če so potrebni popravki

OK HEAD: PersonX acts quickly (izvirnik)
osebaX deluje hitro (prevod) RELATION: xIntent TAIL: to escape from him (izvirnik)
pobegniti pred njim (prevod)

OK HEAD: PersonX acts quickly (izvirnik)
osebaX deluje hitro (prevod) RELATION: xReact TAIL: swift (izvirnik)
swift (prevod)



Statistična analiza podatkov

Podatki

Ročno preverili 10k primerov iz **testne** množice

- Preverjanje sta opravili dve prevajalki

tip množice		število primerov
Množica za treniranje modela (training set)		1,076,880
Množica za preverjanje poteka treniranja (dev set)		102,024
Testna množica	Avtomatično prevedena besedila (vsa)	152,209
	Avtomatično prevedena besedila (A10k)	10,354
	Ročno preverjena besedila (R10k)	10,354

Skupna velikost enaka originalni ATOMIC₂₀ množici podatkov



Slovenski prevodi so javno dostopni na clarin.si – SloATOMIC 2020

Prednosti in slabosti postopka

Podatki

Prednosti

- Relativno enostaven način priprave podatkov za slovenščino
- Ni potrebno ročno nabiranje podatkov (izogib sistemom tipa AWS Mechanical Turk)

Slabosti

- Zanašanje na kvaliteto avtomatskih prevajalnikov
- Ročno preverjanje je časovno zahtevno
 - Morda bi bilo bolje da bi prevajalci sami pripravili prevode
- Primeri ne nujno ujamejo slovenskega konteksta

Razvoj modela za zdravorazumsko sklepanje za slovenščino

Grobelnik M Adrian, Novak Erik, Mladenić Dunja, Grobelnik Marko, 2022. Slomet – Slovenian Commonsense Description. Proceedings of the Slovenian Knowledge Discovery and Databases Conference.

Grobelnik M Adrian, Mladenić Dunja, Grobelnik Marko, 2020. MultiCOMET – Multilingual Commonsense Description. Proceedings of the Slovenian Knowledge Discovery and Databases Conference.

Fino nastavljanje modelov

Evalvacija

- Fino nastavljanje slovenskih GPT modelov z uporabo množice SloATOMIC 2020
- Predikcija modela v obliki:
“{HEAD} {relacija} [GEN]” → “{HEAD} {relacija} [GEN] {TAIL} [EOS]”

HEAD	relacija	TAIL
OsebaX je pomotoma polila	xWant	očistiti oblačila

Vhodno besedilo OsebaX je pomotoma polila xWant [GEN]

Izhodno besedilo OsebaX je pomotoma polila xWant [GEN] očistiti oblačila [EOS]



Izračun uspešnosti modela

Evalvacija

- Podatki imajo za dani par (HEAD, relacija) več različnih TAIL vrednosti
- Izračunamo predikcijo modela za par in primerjamo z dejanskimi vrednostmi v podatkih

Primer iz podatkov

OsebaX je pomotoma polila xWant [GEN] očistiti oblačila [EOS]

OsebaX je pomotoma polila xWant [GEN] pobrisati površino [EOS]

Generiran primer

OsebaX je pomotoma polila xWant [GEN] preobleči oblačila [EOS]



Metrike za analiziranje uspešnosti modela

Evalvacija

metrika	opis metrike
BLEU	Primerjanje in štetje n zaporednih besed ki se pojavijo v napovedanim besedilom in željenim besedilom
CIDEr	Računanje razdalje med napovedanim besedilom in željenim besedilom na podlagi njihovih TF-IDF vektorskih oblik izračunanih z uporabo korenskih oblik njihovih besed
METEOR	Primerjava besed, njenih korenskih oblik in pomena med napovedanim besedilom in željenim besedilom
ROUGE-L	Merjenje najdaljšega skupnega podniza med napovedanim besedilom in željenim besedilom

Rezultati – avtomatično prevedeni podatki

Evalvacija

Jezikovni model	Jezik	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDEr	METEOR	ROUGE-L
Milos/slovak-gpt-j-162M	SLK	0.158	0.068	0.033	0.020	0.236	0.106	0.194
EleutherAI/gpt-neo-125M	ENG	0.263	0.128	0.051	0.023	0.431	0.184	0.351
macedonizer/sl-gpt2	SLV	0.322	0.163	0.087	0.052	0.476	0.220	0.395
gpt-janezek	SLV	0.301	0.148	0.082	0.053	0.490	0.210	0.388
gpt-janez	SLV	0.324	0.174	0.108	0.076	0.508	0.225	0.397
cjvt/gpt-sl-base	SLV	0.338	0.192	0.121	0.086	0.526	0.235	0.413
<i>Originalni model</i>								
COMET(GPT2-XL)	ENG	0.407	0.248	0.171	0.124	0.653	0.292	0.485



Od slovenskih modelov najbolje deluje tisti natreniran z **cjvt/gpt-sl-base**: model javno dostopen na clarin.si – SloMET-ATOMIC 2020

Rezultati – ročno preverjeni podatki

Evalvacija

A10k: Avtomatsko prevedenih 10k primerov

R10k: Ročno preverjenih 10k primerov

Jezikovni model	Testna množica	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDEr	METEOR	ROUGE-L
Milos/slovak-gpt-j-162M	A10k	0.321	0.157	0.079	0.031	0.541	0.217	0.406
	R10k	0.273	0.121	0.068	0.028	0.504	0.189	0.354
EleutherAI/gpt-neo-125M	A10k	0.256	0.069	0.000	0.000	0.498	0.178	0.355
	R10k	0.221	0.046	0.001	0.000	0.472	0.160	0.320
macedonizer/sl-gpt2	A10k	0.327	0.169	0.099	0.064	0.535	0.219	0.408
	R10k	0.266	0.099	0.047	0.023	0.488	0.183	0.350
gpt-janezek	A10k	0.331	0.151	0.078	0.047	0.553	0.214	0.402
	R10k	0.266	0.094	0.041	0.020	0.488	0.181	0.345
gpt-janez	A10k	0.340	0.175	0.101	0.063	0.551	0.229	0.420
	R10k	0.283	0.136	0.089	0.061	0.505	0.194	0.358
cjvt/gpt-sl-base	A10k	0.345	0.193	0.128	0.092	0.568	0.230	0.423
	R10k	0.282	0.149	0.095	0.000	0.509	0.193	0.357



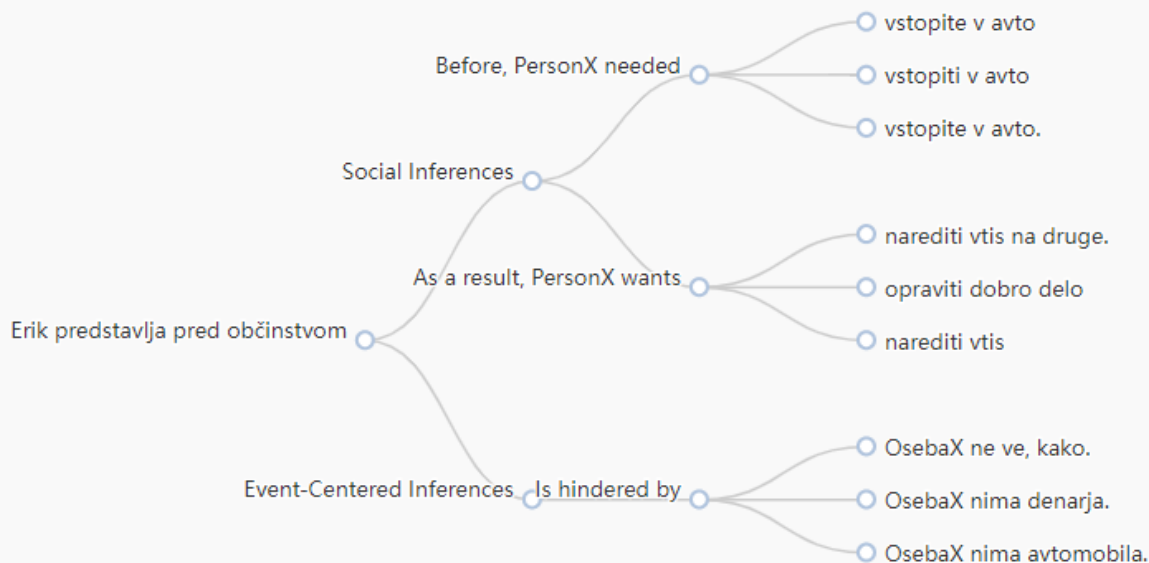
Praktična uporaba modelov

Slovene ▾ People & Events ▾ Erik predstavlja pred občinstvom Submit

Select 1 to 4 descriptors:

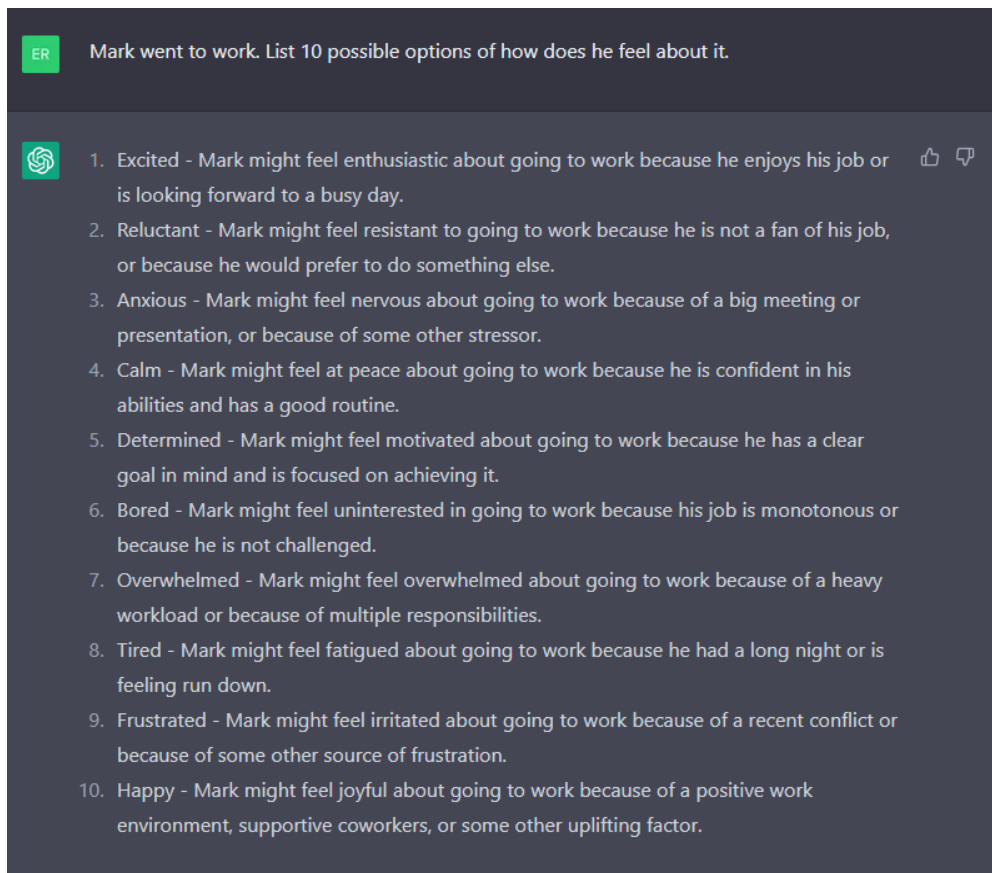
- Before, PersonX needed PersonX is seen as As a result, PersonX feels PersonX then
- As a result, PersonX wants Because PersonX wanted As a result, others feel Others then
- As a result, others want Happens after A possible subevent is Happens before Is hindered by Causes
- As a result, PersonX reasons If there is a blank, it can be filled by

Poskusi: OsebaX hitro reagira, Janez je zelo pomemben, Mojca je pojedla odličen sendvič




Naslednji koraki

- Treniranje modelov na drugih generativnih modelih ki podpirajo slovenščino (BART, GPT-3)
- Izboljšanje prevodov SloATOMIC 2020
- Uporaba ChatGPT za pripravo podatkovne množice za zdravorazumsko sklepanje
 - Analiza vpliva kontekstov
 - poklic, države, karakteristike posameznikov, itd.



ER Mark went to work. List 10 possible options of how does he feel about it.

 1. Excited - Mark might feel enthusiastic about going to work because he enjoys his job or is looking forward to a busy day.

2. Reluctant - Mark might feel resistant to going to work because he is not a fan of his job, or because he would prefer to do something else.

3. Anxious - Mark might feel nervous about going to work because of a big meeting or presentation, or because of some other stressor.

4. Calm - Mark might feel at peace about going to work because he is confident in his abilities and has a good routine.

5. Determined - Mark might feel motivated about going to work because he has a clear goal in mind and is focused on achieving it.

6. Bored - Mark might feel uninterested in going to work because his job is monotonous or because he is not challenged.

7. Overwhelmed - Mark might feel overwhelmed about going to work because of a heavy workload or because of multiple responsibilities.

8. Tired - Mark might feel fatigued about going to work because he had a long night or is feeling run down.

9. Frustrated - Mark might feel irritated about going to work because of a recent conflict or because of some other source of frustration.

10. Happy - Mark might feel joyful about going to work because of a positive work environment, supportive coworkers, or some other uplifting factor.



Razvoj slovenščine v digitalnem okolju

Razvoj jezikovnih tehnologij je ključnega pomena za preživetje jezika v digitalni dobi.

Trajanje: Maj 2022 – Februar 2023

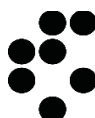
Cilj projekta je zadovoljiti potrebe po **računalniških izdelkih in storitvah s področja jezikovnih tehnologij za slovenski jezik** za raziskovalne organizacije, za podjetja in za širšo javnost. Končni izdelki bodo na uporabniško prijazen način pomagali pri sporazumevanju, sodelovanju, poslovanju, izmenjavi znanja in udeleževanju v družabnih in političnih razpravah, in prispevali k premagovanju jezikovnih meja.



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA KULTURO



EVROPSKA UNIJA
EVROPSKI SKLAD ZA
REGIONALNI RAZVOJ
NALOŽBA V VAŠO PRIHODNOST



Institut
Jožef Stefan

JOTA, 07.03.2023

32

Odsek za
Umetno Inteligenco



Zaključek

Predstavili smo postopek za izdelavo modela za zdravorazumsko sklepanje v slovenščini.



Model za zdravorazumsko sklepanje v slovenščini
SloMET-ATOMIC 2020 (dostopen na clarin.si)

Slovenski prevodi množice ATOMIC₂₀²⁰
SloATOMIC 2020 (dostopen na clarin.si)



Delo opravljeno v sodelovanju z osebami: Slavko Žitnik, Simon Krek, Adrian M Grobelnik, Dunja Mladenčić, Marko Grobelnik, Aljaž Košmerlj, Matej Ulčar, Kristijan Žagar, ...