

Korištenje velikih jezičnih modela kao izvora medicinskih informacija za pacijente i liječnike

Sallabi, Adrian

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, School of Medicine / Sveučilište u Zagrebu, Medicinski fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:105:376847>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom](#).

Download date / Datum preuzimanja: **2024-11-13**



Repository / Repozitorij:

[Dr Med - University of Zagreb School of Medicine Digital Repository](#)



SVEUČILIŠTE U ZAGREBU
MEDICINSKI FAKULTET

Adrian Sallabi

Korištenje velikih jezičnih modela kao izvora medicinskih informacija za
pacijente i liječnike



Zagreb, 2024

Ovaj diplomski rad je izrađen na katedri za Epidemiologiju, medicinsku informatiku i medicinsku statistiku Medicinskog fakulteta Sveučilišta u Zagrebu pod vodstvom doc. dr. sc. Andrije Štajduhara i predan na ocjenu u akademskoj godini 2023/2024.

SADRŽAJ RADA

1. UVOD	1
2. POVIJEST RAZVOJA I DEFINICIJA	2
3. OGRANIČENJA VELIKIH JEZIČNIH MODELA.....	5
4. VELIKI JEZIČNI MODELI U MEDICINI.....	8
5. PRAKTIČNI PRIMJER.....	21
6. ZAKLJUČAK.....	25
7. ZAHVALA	26
8. LITERATURA	27
9. ŽIVOTOPIS	32

SAŽETAK

Adrian Sallabi

Korištenje velikih jezičnih modela kao izvora informacija za pacijente i liječnike

Veliki jezični modeli (LLM) predstavljaju najsuvremenije sustave za obradu prirodnog jezika, sposobne za rješavanje brojnih složenih zadataka u različitim industrijama, uključujući medicinu. Njihove potencijalne primjene u području zdravstva obuhvaćaju kliničku podršku, dijagnostiku te personaliziranu skrb. Međutim, prije njihove šire primjene u zdravstvenom sustavu, potrebno je temeljito razumjeti njihove prednosti i nedostatke. U ovom radu opisane su prednosti i ograničenja velikih jezičnih modela u medicinskom kontekstu, s naglaskom na njihovu pouzdanost pri generiranju medicinskih informacija. Prikazana je povijest razvoja ovih modela, te su detaljno obrađene mogućnosti njihove integracije u kliničku praksu. Razmotreni su etički izazovi poput zaštite privatnosti pacijenata, sigurnosti podataka i problematike halucinacija. Naposljetku, primjena velikih jezičnih modela prikazana je kroz praktični primjer izrade AI agenta temeljenog na GPT-4 modelu, koji je uspješno demonstrirao sposobnost rješavanja specifičnih medicinskih zadataka. Evaluirani su rezultati i objašnjene potencijalne implikacije šire primjene velikih jezičnih modela u zdravstvenoj praksi.

Ključne riječi: medicinske informacije, veliki jezični modeli, GPT-4, klinička praksa

SUMMARY

Adrian Sallabi

Using large language models as a source of medical information for patients and doctors

Large language models (LLMs) represent state-of-the-art systems for natural language processing, capable of solving numerous complex tasks across various industries, including medicine. Their potential applications in healthcare include clinical support, diagnostics, and personalized care. However, before their broader implementation in the healthcare system, it is crucial to thoroughly understand their advantages and limitations. This paper outlines the benefits and drawbacks of large language models in a medical context, with a focus on their reliability in generating medical information. The development history of these models is presented, and the possibilities for their integration into clinical practice are explored in detail. Ethical challenges, such as patient privacy protection, data security, and the issue of hallucinations, are also discussed. Finally, the application of large language models is demonstrated through the example of an AI agent based on the GPT-4 model, which successfully showcased its ability to solve specific medical tasks. The results were evaluated, and the potential implications of broader use of LLMs in medical practice were explained.

Keywords: medical information, large language models, GPT-4, clinical practice

1. UVOD

Veliki jezični modeli (eng. Large language models, LLM) su najsuvremeniji sustavi za obradu prirodnog jezika (eng. Natural language processing) koji su se pokazali sposobni za rješavanje širokog spektra složenih zadataka. Ta sposobnost je otvorila mogućnost za njihovu primjenu u brojnim industrijama, uključujući i medicinsku. Također je dovela do ubrzane popularizacije ovih modela. Unutar samo dva mjeseca od svog lansiranja u studenom 2022. godine, ChatGPT (eng. Chat Generative Pre-trained Transformer) tvrtke OpenAI je premašio čak 100 milijuna korisnika. Korištenje velikih jezičnih modela u svakodnevnom radu postaje sve učestalije, zato je potrebna pažljiva analiza ovih sustava kako bi ih se bolje razumjelo, optimalno iskoristilo i postavilo razumna očekivanja. Prije primjene velikih jezičnih modela u bilo kojem području medicinske prakse važno je poznavati ograničenja ovih sustava i njihove potencijalne opasnosti kako bi se zaštitili u prvom redu pacijenti, a onda i liječnici. Stoga je cilj ovog rada pružanje uvida u velike jezične modele, njihove prednosti i nedostatke, a naročito u kontekstu medicine. Također, kroz pregled dosadašnjih istraživanja, istražiti pouzdanost ovih sustava da generiraju ispravne zdravstvene informacije koje će biti korisne liječnicima i pacijentima. Naposljetku, izradom usmjerenog AI agenta za odgovaranje na pitanja pacijenata i liječnika temeljenog na velikom jezičnom modelu GPT-4, kroz praktični primjer prikazati sposobnost velikih jezičnih modela za rješavanje medicinskih zadataka.

2. POVIJEST RAZVOJA I DEFINICIJA

U ovom poglavlju donosi se kratki pregled složene teme razvoja velikih jezičnih modela koji su rezultat desetljeća znanstvenih istraživanja i tehnoloških napredaka. Veliki jezični modeli su veliki, prethodno trenirani statistički modeli jezika bazirani na modelu strojnog učenja zvanog neuronska mreža [1]. Temeljeni su na području istraživanja pod nazivom modeliranje jezika koje datira još iz 1950ih godina prošlog stoljeća, kada je Shannon primijenio teoriju informacija na ljudski jezik, gdje je mjerio koliko dobro jednostavni n-gram modeli jezika predviđaju ili sažimaju prirodni jezični tekst [1].

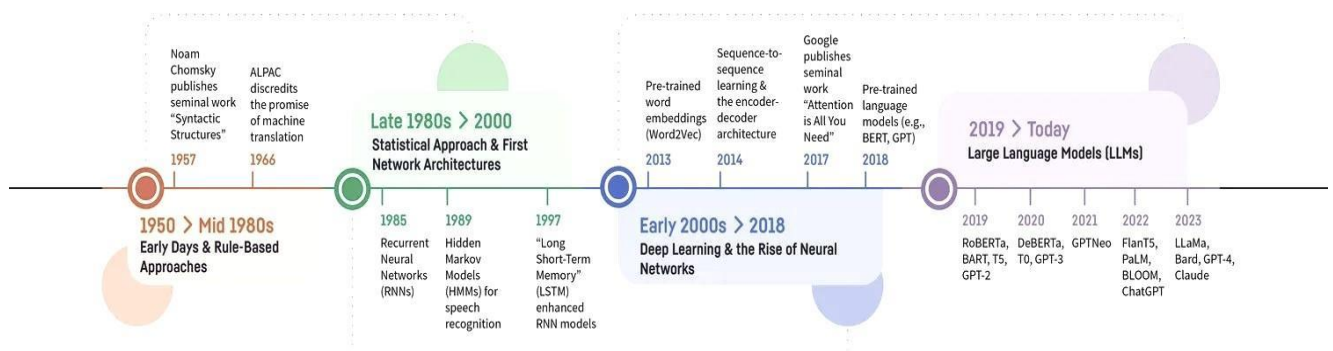
Evolucija tehnologije koja je dovela do razvoja velikih jezičnih modela započinje razvojem statističkih modela jezika. Statistički modeli jezika analiziraju tekst kao niz riječi i procjenjuju vjerojatnost pojave određenog teksta kao produkt vjerojatnosti pojave pojedinačnih riječi. Dominantni oblik statističkog modela jezika su Markovljevi lanci poznati kao n-gram modeli odnosno statistički modeli za predviđanje sljedeće stavke u sekvenci (rečenici) na temelju njezine prethodne riječi [1]. Problematika statističkih modela jezika leži u rijetkosti podataka (eng. Data sparsity) na temelju kojih model procjenjuje vjerojatnost pojave određenih riječi ili sekvenci riječi. Drugim riječima, ovi modeli nisu sposobni u potpunosti oponašati prirodni jezik radi svojih ograničenja, što je dovelo do razvoja ranih neuronskih modela jezika. Rani neuronski modeli jezika rješavaju problem rijetkosti podataka tehnikom ugradbenih vektora (eng. Embedding vectors) [2]. Umjesto predstavljanja svake riječi kao jedinstvenog identifikatora, riječi su mapirane na kontinuirane vektore niske dimenzije, radi čega model bolje razumije i predviđa riječi u tekstu [3]. Rani neuronski modeli jezika bili su trenirani na specifičnim podacima relevantnim za određeni zadatak poput pretraživanja ili prevođenja, za razliku od prethodno treniranih jezičnih modela koji su uslijedili nakon njih. Prethodno trenirani jezični modeli donose značajan napredak u području obrade prirodnog jezika, što je dovelo do promjene standarda strojnog učenja jezičnih modela s nadziranog učenja na prethodno treniranje praćeno finim podešavanjem [4,5].

Veliki napredak u obradi prirodnog jezika predstavljaju transformeri, modeli koji su prvi put predstavljeni 2017. godine u članku "Attention is All You Need" tima istraživača iz Googlea [6]. Prije razvoja transformera, neuronski modeli za obradu prirodnog jezika oslanjali su se na složene rekurentne i konvolucijske neuronske mreže koje su unatoč tome što su bile donekle korisne imale brojna ograničenja [6]. U odnosu na njih, transformeri predstavljaju jednostavnije mrežne arhitekture temeljene na mehanizmima pažnje (eng. Attention mechanism). U znanstvenom članku, Vaswani i suradnici prikazuju kako su transformatori na dva zadatka strojnog prevođenja pokazali superiornost u kvaliteti te istodobno zahtijevali znatno manje vremena za treniranje [6]. Naposljetku dolazimo do danas široko poznatih velikih jezičnih

modela. Veliki jezični modeli uglavnom se odnose na neuronske modele jezika bazirane na transformerima koji sadrže desetke do stotine milijardi parametara, a prethodno su trenirani na ogromnim količinama tekstualnih podataka, kao što su PaLM (eng. Pathways language model), LLaMA (eng. Large language model Meta AI) i GPT-4 (eng. Generative pre-trained transformer 4) [1]. Veliki jezični modeli prolaze kroz tri osnovne faze treniranja: prethodno treniranje, nadzirano fino podešavanje (eng. Supervised fine tuning, SFT) i učenje podrškom s povratnim informacijama od ljudi (eng. Reinforcement learning from human feedback, RLHF). U usporedbi s prethodno treniranim jezičnim modelima, veliki jezični modeli su puno veći u smislu broja parametara te pokazuju snažnije sposobnosti razumijevanja i generiranja jezika. Također, veliki jezični modeli u usporedbi s modelima jezika manjeg opsega imaju takozvane pojavne sposobnosti poput učenja u kontekstu, gdje uče novi zadatak iz malog skupa primjera predstavljenih u zadanom upitu. Sposobnost slijeđenja uputa, gdje modeli nakon prilagođavanja instrukcijama, mogu slijediti upute za nove vrste zadataka bez korištenja eksplicitnih primjera i sposobnost višekratnog zaključivanja koja im omogućuje riješiti složeni zadatak razbijanjem tog zadatka na međukorake zaključivanja [1]. Integracijom velikih jezičnih modela specijaliziranih za obradu prirodnog jezika s drugim modelima, poput prethodno treniranih temeljnih modela, koji obrađuju različite oblike podataka, dobiveni su suvremeni sustavi umjetne inteligencije sposobni za rješavanje širokog spektra složenih zadataka. Primjeri takvih sustava uključuju ChatGPT tvrtke OpenAI, Googleov Gemini i druge.

2.1. AI AGENTI

Veliki jezični modeli danas se mogu koristiti kao agenti umjetne inteligencije. Ova upotreba uključuje primjenu dizajniranja upita (eng. Prompt engineering) za prilagodbu modela specifičnim područjima, čime se pokušava povećati njihova iskoristivost. Na taj način moguće je izraditi specijaliziranog agenta za područje medicine koji rješava određene zadatke prema unaprijed definiranim smjernicama upisanim u njegov sustav (eng. System prompt). Na primjer, moguće je izraditi agenta koji služi za odgovaranje na pitanja pacijenata ili agenta koji pomaže kliničaru u dijagnostici i terapijskom pristupu.



Slika 1. Povijest modeliranja prirodnog jezika, izvor: preuzeto sa stranice medium.com

3. OGRANIČENJA VELIKIH JEZIČNIH MODELA

Kapacitet velikih jezičnih modela da rješavaju složene zadatke i pritom pokazuju sposobnosti koje omogućuju njihovu široku upotrebu. No, kako bi se veliki jezični modeli sigurno primijenili, naročito u osjetljivom području medicine, važno je uzeti u obzir njihova ograničenja. U ovom poglavlju će ukratko biti obrađena ograničenja velikih jezičnih modela s naglaskom na problematiku generiranja besmislenog sadržaja zvanog halucinacija.

3.1. DEFINICIJA I KLASIFIKACIJA HALUCINACIJA

U kontekstu velikih jezičnih modela halucinacije se odnose na pogrešne izlaze (eng. output) koje generiraju veliki jezični modeli, drugim riječima daju odgovore koji izgledaju uvjerljivo, ali nisu utemeljeni na stvarnim podacima, besmisleni su ili netočni. Halucinacije se mogu podijeliti prema različitim kriterijima, ovisno o prirodi, uzrocima i manifestacijama tih pogrešaka. Razlikujemo dva načina klasifikacije halucinacija: fenomenološku i mehaničku. Fenomenološka klasifikacija dijeli halucinacije prema posljedicama, dok se mehanička klasifikacija odnosi na unutarnje mehanizme i uzroke pojave halucinacija. [7,8]. U nastavku se daje kratki osvrt na fenomenološku klasifikaciju halucinacija.

Konvencionalno razlikujemo dvije glavne vrste halucinacija; unutarnje i vanjske. Unutarnje halucinacije se odnose na izlaze koji proturječe danom unosu ili upitu (eng. prompt), odnosno model pogrešno koristi ili tumači informacije iz danog upita [8]. Na primjer, izlaz koji generira model: „Alexander Fleming je 1930. godine otkrio djelovanje penicilina kao antibiotika na uzorku stafilokokne bakterije“. Proturječi ispravnom upitu korisnika: „Alexander Fleming je 1928. godine otkrio direktno djelovanje penicilina kao antibiotika na uzorku stafilokokne bakterije“.

Vanjske halucinacije nastaju kad model generira informacije koje ne mogu biti verificirane ulaznim podacima odnosno na temelju upita ne možemo znati jesu li informacije u danom izlazu ispravne ili neispravne. Drugim riječima, model generira tekst s vanjskim informacijama koje se ne nalaze u upitu korisnika. Vanjska halucinacija ne mora biti pogrešna jer vanjske informacije u izlazu mogu biti ispravne. Takve, činjenično ispravne halucinacije mogu biti korisne jer predstavljaju dodatno znanje i kontekst u generiranom sadržaju. Unatoč tome, i takvim informacijama je potrebno pristupiti s oprezom jer nije poznat njihov izvor i kontekst [8,9].

3.2. UZROCI HALUCINACIJA

Za razliku od fenomenološke, mehanička klasifikacija se temelji na identifikaciji unutarnjih mehanizmima i uzroka koji dovode do halucinacija. Osnovni uzroci halucinacije se općenito povezuju s problemima podataka, treniranja i zaključivanja.

3.2.1. Halucinacije uzrokovane podacima

Halucinacije povezane s podacima u velikim jezičnim modelima prvenstveno su posljedica pogrešnih izvora podataka i inferiorne upotrebe podataka na kojima su modeli trenirani. Pogrešne informacije i pristranosti izvora podataka mogu dovesti do generiranja pristranih izlaza. Na primjer, neke zablude poput tvrdnje „Thomas Edison je izumio žarulju“ su široko rasprostranjene, ukoliko je model prethodno treniran na takvim pogrešnim podacima ponavljat će ih pritom generirajući netočne izlaze [9]. Takvi izlazi se nazivaju imitativnim (ponovljenim) netočnostima.

3.2.2. Halucinacije uzrokovane treniranjem

Prilikom treniranja, veliki jezični modeli prolaze različite faze poput temeljnog prethodnog treniranja gdje modeli uče opće reprezentacije te preuzimaju široko opće znanje i naknadnog usklađivanja gdje su modeli usklađeni kako bi se bolje prilagodili potrebama korisnika. Svaka od tih faza predstavlja jedinstvene izazove koji mogu uzrokovati halucinacije [9].

3.2.3. Halucinacije uzrokovane zaključivanjem

Zaključivanje se odnosi na proces generiranja odgovora ili zaključaka na osnovu zadanog ulaznog teksta. Kada korisnik postavi pitanje ili zada neki zadatak, model koristi svoje prethodno naučene obrasce i statističke informacije kako bi proizveo zaključak o tome što bi mogao biti odgovarajući odgovor. Jedan od uzroka takvih halucinacija su ograničenja vezana uz korak koji slijedi nakon temeljnog prethodnog treniranja i usklađivanja, tzv. dekodiranje. Dekodiranje se odnosi na fazu generiranja teksta nakon obrade ulaznih podataka (upita). Određena ograničenja u ovom koraku poput nasumičnosti i nesavršene reprezentacije podataka mogu dovesti do halucinacija zaključkom [9].

3.3. UBLAŽAVANJE HALUCINACIJA

Da bi se fenomen halucinacija ublažio potrebno je identificirati temeljne uzroke njihove pojave te razviti strategije za njihovo rješavanje. Trenutna rješenja za ublažavanje halucinacija kod velikih jezičnih modela mogu se kategorizirati u tri skupine: ispravljanje tijekom treniranja, ispravljanje tijekom generiranja i ispravljanje uz pomoć dohvaćanja podataka, u skladu s podjelom halucinacija prema uzroku [10].

3.4. OSTALA OGRANIČENJA

Važno je naglasiti kako su veliki jezični modeli statistički sustavi koji prvenstveno služe predikciji sljedećeg tokena (osnovna jedinica teksta koju model obrađuje) u rečenici. Iz tog razloga sadrže unutarnja ograničenja povezana sa temeljnim principima njihovog rada. Veliki jezični modeli sami po sebi ne posjeduju memoriju, odnosno ne mogu zapamtiti čak ni ono što im je poslano u prethodnom upitu. To je važno ograničenje za mnoge slučajeve upotrebe koji zahtijevaju neki oblik memorije. Nadalje, oni su probabilistički, ukoliko se modelu pošalje isti upit nekoliko puta, najčešće će se dobiti različiti izlazi. Iako postoje parametri poput temperature, koji mogu ograničiti varijabilnost u odgovorima, to je unutarnje svojstvo njihovog treniranja koje može biti problematično. Jedan od važnih problema je i nedostatak ažuriranja informacija, naime veliki jezični modeli sami po sebi nemaju pristup osvježnim podacima. To je važno naročito u područjima poput medicine u kojima se informacije učestalo mijenjaju. Veliki jezični model sam po sebi ne zna trenutni datum ili vrijeme i nema pristup bilo kojoj informaciji koja nije bila prisutna u njegovom skupu za treniranje. Uz to, veliki jezični modeli su složeni sustavi koji troše puno resursa. To znači da je potrebno mnogo sofisticiranih strojeva za njihovo treniranje i posluživanje [1]. Dodatna važna ograničenja u području medicine bit će opisana zasebno u odgovarajućem poglavlju.

4. VELIKI JEZIČNI MODELI U MEDICINI

4.1. PRIMJENE

Veliki jezični modeli, naročito integrirani kao dio multimodalnih sustava umjetne inteligencije koji kombiniraju tekst, slike i druge oblike podataka, posjeduju sposobnosti koje se potencijalno mogu primijeniti na širok spektar medicinskih zadataka. Njihove sposobnosti nisu ograničene samo na generiranje teksta, već su prilagođene i za složene obrade podataka, analize obrazaca i donošenje odluka, što je osobito važno u medicinskom kontekstu. Međutim, kako bi se takvi sustavi uistinu koristili u praksi, ključno je razumjeti njihove nedostatke i etičke nedoumice povezane s njihovom primjenom. U nastavku se razmatraju primjene velikih jezičnih modela u različitim područjima medicine, s posebnim osvrtom na njihove koristi i ograničenja.

4.1.1. MEDICINSKI ZADACI

Zadatke strojnog učenja u kontekstu medicine možemo podijeliti na zadatke razlikovanja (eng. Discriminative tasks) i generativne zadatke (eng. Generative tasks) [10]. Zadaci razlikovanja se odnose na zadatke koji uključuju kategorizaciju i svrstavanje podataka u specifične kategorije na temelju danih ulaznih podataka. U području medicine ulazi koje veliki jezični modeli rješavaju kao zadatke razlikovanja mogu biti medicinska pitanja, kliničke bilješke, dokumenti, istraživački radovi i elektronički zdravstveni zapisi pacijenata. Tako veliki jezični modeli na temelju medicinske literature, medicinskog znanja i podataka o pacijentima mogu pružiti odgovore na klinička pitanja pacijenata ili liječnika. Na primjer, pacijent može postaviti upit o svojim simptomima, dijagnozi i terapiji. Takva upotreba može biti od pomoći pacijentima, a i liječnicima povećanjem učinkovitosti postavljanja dijagnoze. Također, modeli imaju sposobnost automatski identificirati i kategorizirati ključne informacije poput simptoma, lijekova, dijagnoza i laboratorijskih rezultata pacijenata, čime se pomaže u organiziranju i upravljanju podacima pacijenata. Za razliku od zadataka razlikovanja, generativni zadaci se odnose na zadatke strojnog učenja koji zahtijevaju generiranje ispravnog novog sadržaja na temelju vanjskih informacija koje se ne nalaze u ulazu odnosno upitu. Ti zadaci uključuju sažimanje, generiranje i pojednostavljenje medicinskog teksta [10]. Sažimanje dugog i kompliciranog medicinskog teksta se u kliničkoj praksi može koristiti kako bi se gotovo trenutno generirala skraćena verzija teksta pri čemu se naglašavaju najvažniji dijelovi bez gubitka osnovne medicinske informacije. Sažeti se mogu medicinski izvještaji, otpusna pisma, radiološki nalazi i slični dokumenti čime se olakšava pristup pacijenata i liječnika važnim zdravstvenim podacima. Pojednostavljenje medicinskog teksta se odnosi na sposobnost

modela da pojednostavljaju složeni medicinski žargon i objasne medicinske pojmove kako bi bili lakše razumljivi široj populaciji. Ulazi za funkciju generiranja medicinskog teksta mogu uključivati medicinsku dokumentaciju koja sadrži nalaze pretraga, dijagnoze i provedenu terapiju, a na temelju kojih model generira izlaz koji može sadržavati zdravstvene savjete, upute za pacijenta, medicinske dijagnoze i slično.

4.1.2. KLINIČKE PRIMJENE

Veliki jezični modeli imaju potencijal za pružanje kliničke podrške liječnicima kroz omogućavanje brzog i preciznog pristupa relevantnim informacijama. Rješavanjem generativnih zadataka poput sažimanja i pojednostavlivanja teksta mogu povećati učinkovitost administrativnih poslova u zdravstvu. Primjerice, automatizacija procesa kao što je generiranje medicinskih izvještaja može rasteretiti zdravstveno osoblje od repetitivnih i vremenski zahtjevnih zadataka, omogućujući im da se više posvete neposrednoj skrbi za pacijente. Osim toga, veliki jezični modeli mogu analizirati medicinsku literaturu i pružiti personalizirane preporuke prilagođene specifičnim potrebama pacijenata, što može poboljšati kvalitetu zdravstvene skrbi i donošenje odluka. U nastavku slijedi pregled kliničkih primjena velikih jezičnih modela u medicini, s posebnim naglaskom na njihov potencijal za unapređenje dijagnostičkih procesa, podršku kliničkom odlučivanju, automatizaciju administrativnih zadataka te personaliziranu skrb o pacijentima.

4.1.2.1. Donošenje kliničkih odluka

Veliki jezični modeli mogu služiti kao asistenti liječnicima pri donošenju odluka o skrbi za pacijente, pružajući im uvid u znanstvenu literaturu, uključujući najnovije smjernice za liječenje, dijagnostičke metode ili informacije o lijekovima [11]. Za razliku od konvencionalnih alata za pretraživanje, ovi modeli mogu pružiti informacije specifične za pojedinog pacijenta, čime se povećava učinkovitost pretraživanja kliničkih podataka. Na taj način kliničarima omogućuju brže i preciznije dobivanje ciljanih odgovora koji im pomažu u donošenju informiranih odluka. Jedno presječno istraživanje u kojemu su liječnici ocjenjivali odgovore ChatGPT-4 modela na pitanja o dijagnozi i planu liječenja za otorinolaringološke pacijente je pokazalo visoku suglasnost dijagnoza i plana liječenja između liječnika i ChatGPT-a [11,12].

4.1.2.2. Dijagnostika

Jedna od najperspektivnijih primjena velikih jezičnih modela je ona u domeni dijagnostike. Modeli potencijalno mogu doprinijeti učinkovitom postavljanju dijagnoze prvenstveno proširivanjem diferencijalne dijagnoze na temelju unesene medicinske dokumentacije pacijenta. Modeli su sposobni temeljito analizirati različite kliničke podatke, poput simptoma, povijesti bolesti i rezultata dijagnostičkih testova, što im omogućuje brzo generiranje mogućih dijagnoza. Također ih je moguće integrirati s različitim bodovnim sustavima koji se koriste za procjenu stadija bolesti poput određivanja TNM stadija, Ransonovih kriterija za pankreatitis i drugih dijagnostičkih kriterija [11]. U kontekstu dijagnostike, zanimljivo je istraživanje u kojemu su liječnici procjenjivali diferencijalne dijagnoze koje su generirali veliki jezični modeli GPT-4 i GPT-3.5 za 75 kliničkih slučajeva. Studija je pokazala kako je model GPT-4 generirao ispravnu diferencijalnu dijagnozu u otprilike dvije trećine slučajeva, ali oba modela su najčešće pogriješila u odabiru najvjerojatnije dijagnoze [13]. Ovaj primjer pokazuje kako je modele za sad moguće koristiti isključivo kao asistente kliničarima u izvođenju medicinskih zadataka poput proširivanja diferencijalne dijagnoze ili davanja preporuka o liječenju. Ključno je da kliničari pristupaju ovim modelima s oprezom, svjesni njihovih nedostataka i koristeći ih kao dodatni alat, a ne zamjenu za svoje stručno znanje.

4.1.2.3. Klinička izvješća

Značajan potencijal modela je u automatizaciji i poboljšanju pisanja kliničkih izvještaja, pritom smanjujući administrativno opterećenje liječnika. Njihova upotreba bi omogućila brže i preciznije generiranje standardizirane dokumentacije, kao što su radiološki izvještaji, otpusna pisma ili klinička pisma pacijentima. S obzirom da dokumentacija i administrativni zahtjevi zauzimaju oko 25% radnih dana kliničara, pojednostavljenje procesa pisanja dokumentacije može smanjiti opterećenje zdravstvenih djelatnika i pomoći u ublažavanju sindroma sagorijevanja [14, 15]. Jedna od sposobnosti velikih jezičnih modela integriranih u multimodalne sustave koji mogu obrađivati i druge oblike podataka je automatizirana sinteza medicinskih izvještaja iz slikovnih ili glasovnih podataka [16]. Ovi modeli analiziraju medicinske slike i generiraju bilješke koje se potom koriste kao dodatni unos uz tekstualne upute za velike jezične modele. Kombinacijom vizualnih i tekstualnih podataka, model stvara precizne i strukturirane izvještaje koji odgovaraju unaprijed definiranim parametrima [10]. Jedno istraživanje koristi prethodno trenirani jezični model i učenje u kontekstu za generiranje kliničkih bilješki iz razgovora između liječnika i pacijenta [17]. Ovakvi integrirani sustavi imaju potencijal da značajno unaprijede proces automatizacije generiranja medicinskih izvještaja [16]. Postoji nekoliko specijaliziranih velikih jezičnih modela razvijenih upravo za generiranje

kliničkih izvještaja. Na primjer, modeli kao što su Med-Gemini, LLaVA-Med i Med-Flamingo mogu se koristiti u širokom medicinskom kontekstu, dok su ChatCAD i MAIRA razvijeni specifično za radiološke izvještaje. Ovi modeli su prilagođeni specifičnim medicinskim područjima te radi toga postižu visoku točnost u bilježenju relevantnih informacija. Iako modeli mogu generirati izvještaje koji su detaljni i precizni, suočavaju se s izazovima poput halucinacija i doslovne interpretacije podataka bez razumijevanja kliničkog konteksta. Također, njihovi izvještaji su često su manje sažeti od onih koje pišu ljudi, a objektivna procjena njihove točnosti i kvalitete još uvijek predstavlja izazov [10].

4.1.2.4. Personalizirana skrb za pacijente

Korištenje velikih jezičnih modela u svrhu unapređenja skrbi za pacijente podrazumijeva i sposobnost modela da pruže personalizirane zdravstvene informacije pacijentima. Modeli, za razliku od zdravstvenog osoblja, nemaju vremenska ograničenja te predstavljaju nepresušan izvor informacija o zdravstvenom stanju pacijenta. Kao takvi, ukoliko pružaju točne i korisne zdravstvene informacije mogu pacijentima služiti kao podrška kroz sve faze liječenja. Takav oblik korištenja modela u kliničkoj praksi potencijalno bi mogao poboljšati ishode pacijenata, a preciznim praćenjem pacijenata omogućiti zdravstvenim djelatnicima donošenje odluka na temelju preciznih podataka, što rezultira kvalitetnijom skrbi [15].

4.1.2.5. Mentalna podrška

Još jedna od mogućih primjena velikih jezičnih modela je pružanje terapijske komunikacije pacijentima. Modeli se mogu koristiti kao podrška u očuvanju mentalnog zdravlja pacijenata, pružajući im kontinuiranu skrb bez vremenskih ograničenja. Kod ovakvog oblika primjene do izražaja dolazi i mogućnost smanjenja prepreka poput srama ili nelagode koji često prate komunikaciju s ljudskim terapeutom [10]. Istraživanja su pokazala da pacijenti često iskazuju veću spremnost za razgovor o osjetljivim temama, poput mentalnog zdravlja, s digitalnim sugovornicima, nego s ljudima, što može dovesti do veće otvorenosti i preciznijih dijagnoza [18,19]. Jedna od značajnih prednosti modela u pružanju mentalne skrbi je povećanje dostupnosti zdravstvene skrbi pacijentima koji imaju financijske prepreke. Klasične psihoterapije često su dugotrajne i skupe, što mnogim pacijentima otežava pristup liječenju. Međutim, važno je pristupiti ovim tehnologijama s oprezom. Dok modeli mogu pružiti korisne informacije i osnovnu podršku, oni ne mogu zamijeniti ljudske terapeute kada je riječ o emocionalnom razumijevanju i empatiji, što su ključne komponente u liječenju mentalnih

poremećaja. Također, uvijek postoji određena razina rizika od neprikladnih ili čak štetnih odgovora, osobito ako modeli nisu pravilno trenirani ili nadzirani [10].

4.1.2.6. Prijevod medicinskog jezika

Razlikujemo dva glavna oblika prevođenja medicinskog jezika u kontekstu velikih jezičnih modela: prevođenje medicinske terminologije s jednog jezika na drugi i prevođenje medicinskog dijaloga kako bi bio razumljiv laičkoj populaciji [10, 20, 21]. Svaki od ovih oblika prevođenja je važan za unaprjeđenje medicinske komunikacije ne samo između zdravstvenog osoblja i pacijenata već i drugih skupina. Ova sposobnost modela može premostiti komunikacijski jaz između pružatelja zdravstvenih usluga i pacijenata potičući informirano donošenje odluka, zajedničko razumijevanje i zadovoljstvo pacijenata [10]. Nadalje, pojednostavljenje složenog medicinskog žargona pobliže upućuje pacijenta u njegovo zdravstveno stanje, čime se povećava njegova informiranost, a samim time i voljnost za sudjelovanje u terapijskom procesu. Unatoč ovim prednostima i kod ovakvih zadataka modeli mogu biti skloni pogrešci. Kako bi se smanjila mogućnost netočnih prijevoda, preporučuje se da se modeli treniraju na pouzdanim izvorima kao što su medicinski udžbenici i recenzirani medicinski časopisi [10]. To će pomoći u očuvanju točnosti terminologije i osigurati da prijevod odražavaju pravi klinički kontekst, minimalizirajući rizik od nesporazuma ili dezinformacija koje mogu imati i pravne posljedice.

4.1.2.7. Kirurgija

Veliki jezični modeli pokazuju sposobnosti koje se mogu primijeniti i u kirurškom okruženju. Modeli mogu pomoći u izradi detaljnih operativnih izvještaja i bilješki o napretku pacijenta [11, 22]. Također mogu generirati perioperativne uključujući upute za otpuštanje pacijenta. Dobro napisane otpusne upute su važne jer se pokazalo kako loše napisane upute dovode do lošijih ishoda kod kirurških pacijenata [11, 23]. Kao u ostalim granama medicine i kirurški pacijenti mogu koristiti veliki jezični model kako bi se informirali o svom zdravstvenom stanju. Primjerice, mogu dobiti usmjerene podatke o terapiji, smjernicama kojih se moraju pridržavati prije i nakon kirurškog zahvata što može utjecati na ishode liječenja. Također, mogu podržati kliničko odlučivanje, pomažući u odabiru između kirurških i ne-kirurških intervencija, uz procjenu preoperativnih rizika kako bi se osigurao idealan kandidat za operaciju [11].

4.1.2.8. Medicinska robotika

Medicinska robotika donosi revoluciju u zdravstvo, omogućujući preciznost u različitim aspektima, poput kirurških zahvata i medicinskog snimanja [10]. Veliki jezični modeli pokazuju potencijal za dodatno unapređenje ovih sustava djelujući komplementarno tehnologiji robotike. Pokazalo se kako kirurški roboti potpomognuti velikim jezičnim modelima omogućuju minimalno invazivne zahvate s većom točnošću i kraćim vremenom oporavka pacijenata [10, 24]. Zanimljiv primjer primjene modela njihovom integracijom s robotskim sustavom je sustav SuFIA, prvi okvir za robotske kirurške asistente s poboljšanom spretnosti vođene prirodnim jezikom. SuFIA integrira sposobnosti zaključivanja velikog jezičnog modela GPT-4 turbo s modulima za percepciju kako bi postigla planiranje i kontrolu robotskog sustava u izvršavanju manjih kirurških zadataka [25].

4.1.2.9. Radiologija

U radiologiji, veliki jezični modeli mogu poboljšati analizu medicinskih slika, interpretaciju nalaza i automatizaciju izvještaja. Modeli poput ChatGPT 3.5 i 4.0, Google Bard-a i Microsoft Bings-a pokazali su se kao moćni alati za pojednostavljenje radioloških izvještaja kako bi ih približili pacijentima [26]. U području medicinskog snimanja, integrirani sustavi poput UltrasoundGPT-a koriste modele kako bi unaprijedili zahvate vođene ultrazvukom, koristeći bazu znanja koja omogućuje precizno planiranje pokreta tijekom postupaka. Slično tome, veliki jezični modeli mogu se koristiti za kontrolu robotskih sustava za X-zrake, poput uređaja Brainlab Loop-X koji poboljšava preciznost i učinkovitost kirurških zahvata vođenih X-zrakama [10].

4.1.3. IZVOR MEDICINSKIH INFORMACIJA

4.1.3.1. Medicinski testovi

Veliki jezični modeli su se pokazali izrazito uspješnima u rješavanju pitanja iz ispita namijenjenih studentima medicine. Istraživanje objavljeno u drugom mjesecu 2023. godine uspoređivalo je uspješnost američkih studenata medicine na testu USMLE 1 (eng. US medical licensing exam 1) s uspješnosti danas zastarjele verzije velikog jezičnog modela ChatGPT-3.5. Istraživanje je pokazalo da je ChatGPT prvi model koji je postigao točnost na razini ili blizu praga prolaznosti od 60%. Posebice je zanimljivo da je ChatGPT uspio postići ovaj rezultat bez specijaliziranih uputa [27]. Moderna verzija OpenAI modela velikog jezičnog modela ChatGPT-4 pokazuje daleko impresivnije rezultate na provjerama znanja korištenim za testiranje američkih studenata. Tako GPT-4 prelazi prag prolaznosti na svakom koraku američkog ispita za dobivanje medicinske licence s rezultatima između 80 i 90% riješenosti [15,28]. Osim toga, veliki jezični modeli specifično trenirani na skupovima medicinskih podataka poput modela Med-Palm 2 pokazuju veliku sposobnost za rješavanje složenih medicinskih zadataka [29]. Tim istraživača Google-a je prikazao kako Med-PaLM 2 pokazuje izvrsne rezultate u odgovaranju na medicinska pitanja u formatu višestrukog izbora i dugih odgovora [30]. Iako modeli pokazuju izvrsne rezultate na strukturiranim testovima, mogu li se te sposobnosti koristiti u stvarnim nepredvidivim kliničkim scenarijima i do koje mjere, potrebno je istražiti. Studija koja je evaluirala sposobnosti kliničke racionalizacije ChatGPT-4 modela na temelju odgovora koje je ponudio na USMLE pitanja, prikazala je da GPT-4 pokazuje napredak u razumijevanju složenih medicinskih scenarija, ali i dalje ima izazove u interpretaciji dijagnostičkih testova [31].

4.1.3.2. Odgovaranje na upite pacijenata

Mogu li se modeli koristiti kao izvori medicinskih informacija se može procijeniti i istraživanjima koja se bave ispitivanjem odgovora velikih jezičnih modela na pitanja liječnika i pacijenata. U doba Interneta, pacijenti sve više teže prema samostalnom informiranju o svom zdravstvenom stanju. Međutim, često nemaju potrebna medicinska znanja za razumijevanje i odabir dobivenih informacija. Kao rezultat toga, sve se češće obraćaju liječnicima tražeći odgovore na svoja pitanja putem e-pošte ili drugih oblika digitalne komunikacije. Veliki jezični modeli korišteni kao AI asistenti za odgovaranje na pitanja pacijenata mogli bi predstavljati adekvatno rješenje za zahtjeve pacijenata uz što manje dodatno opterećenje zdravstvenog osoblja. Jedno od ranih istraživanja na ovu temu je ono objavljeno u časopisu JAMA (Journal of the American

Medical Association), koje je uspoređivalo odgovore liječnika s platforme Reddit s odgovorima ChatGPT-3.5. Rezultati su pokazali da su odgovori ChatGPT-3.5, često bili opsežniji i sadržajno bogatiji u usporedbi s onima koje su davali liječnici. Međutim, također se istaknulo da su odgovori ChatGPT-a ponekad bili manje precizni ili su sadržavali nepotpune informacije [32].

4.1.3.3. Odgovaranje na upite liječnika

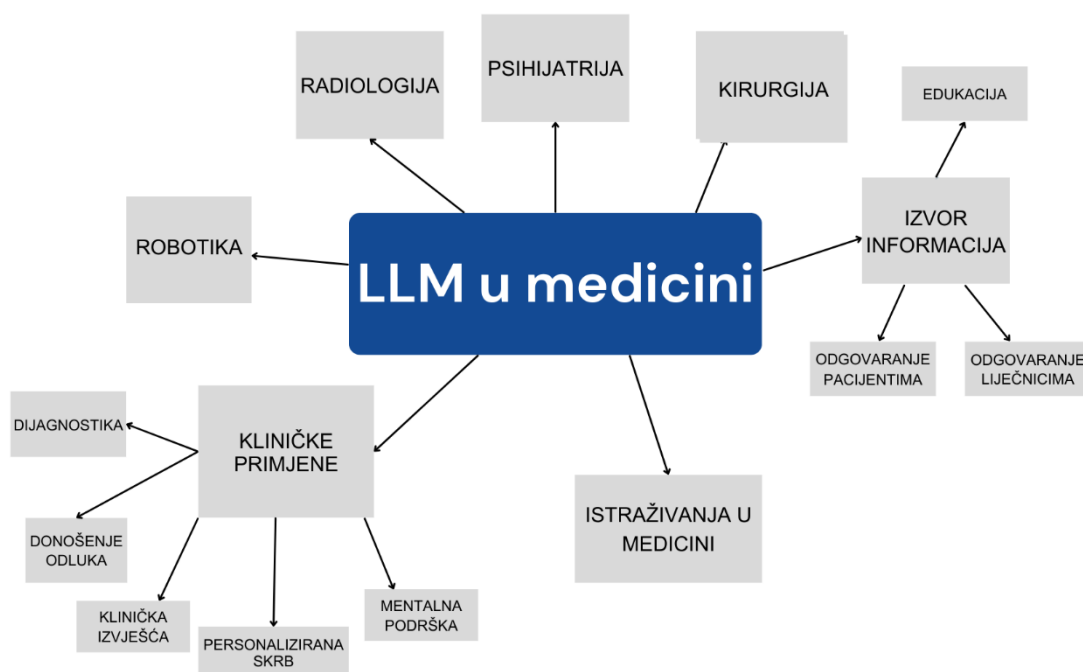
Veliki jezični modeli mogu pomoći liječnicima da se učinkovito i brzo informiraju o medicinski relevantnim temama. Ova tehnologija ima potencijal za unaprjeđenje načina na koji liječnici pristupaju medicinskim informacijama. Za razliku od tradicionalnih metoda pretraživanja podataka, modeli imaju sposobnost generirati usmjerene i personalizirane odgovore na temelju specifičnih kliničkih pitanja, čime se omogućava brži pristup relevantnim informacijama. Uz to informacije koje modeli pružaju, ukoliko je to potrebno, mogu biti sažete i pojednostavljene što liječnicima može uštedjeti vrijeme koje bi inače proveli analizirajući opsežnu literaturu. Međutim, kako bi se liječnici mogli osloniti na velike jezične modele, ključno je osigurati točnost i relevantnost podataka koje generiraju. Istraživanje koje je procjenjivalo točnost i pouzdanost odgovora ChatGPT-a na 284 pitanja postavljena od strane 33 liječnika različitih specijalnosti je pokazalo da ChatGPT generira uglavnom točne informacije na raznolika medicinska pitanja prema ocjeni akademskih liječnika specijalista, iako s važnim ograničenjima [33].

4.1.3.4. Pristup znanstvenim informacijama

Jedan od glavnih izazova u znanstvenom istraživanju jest filtriranje relevantnih informacija iz velikih baza podataka. Veliki jezični modeli imaju kapacitet za ubrzavanje ovog procesa tako što prepoznaju ključne pojmove i brzo generiraju sažetke, čime se medicinski istraživači mogu usmjeriti na bitne aspekte svojih studija. Na primjer, modeli poput ChatGPT-a i Med-PaLM-a sposobni su pretražiti medicinske baze podataka, poput PubMed-a, te korisnicima omogućiti razumijevanje složenih znanstvenih članaka na jednostavniji način. Uz to, modeli mogu pomoći istraživačima u pronalaženju smjernica za liječenje temeljene na dokazima, omogućujući brži pristup novim istraživanjima ili preporukama [15]. Ipak, potrebno je oprezno pristupiti korištenju modela za pristup znanstvenim informacijama zbog mogućnosti dezinformacija ili pristranosti.

4.1.4. MEDICINSKO OBRAZOVANJE

Veliki jezični modeli bi mogli biti korisni za edukaciju studenata zdravstvenih usmjerenja. Modeli se mogu koristiti kao asistenti za učenje, koji pomažu studentima i zdravstvenim profesionalcima razumjeti kompleksne medicinske koncepte. Mogu značajno doprinijeti različitim obrazovnim strategijama pružanjem kvalitetnih sažetaka, prezentacija, prijevoda, objašnjenja i vodiča korak po korak, te nudeći kontekstualizaciju za širok raspon tema, uz mogućnost prilagodbe složenosti i stila odgovora [15]. Jedno istraživanje, izrađeno iz perspektive studenata medicine identificiralo je tri specifična pozitivna primjera upotrebe ChatGPT-a u medicinskoj edukaciji studenata: olakšavanje razmišljanja o diferencijalnoj dijagnozi, pružanje interaktivnih vježbi slučajeva i pomoć pri rješavanju pitanja s višestrukim izborom odgovora. No, istraživači su naglasili i ključna ograničenja modela u medicinskom obrazovanju poput izostanka učenja o razumijevanja osjetilnih i neverbalnih znakova, razvijanja odnosa i interpersonalne interakcije, te usklađivanja s općim ciljevima medicinskog obrazovanja [34]. Korištenje modela kao obrazovnih alata izaziva zabrinutost jer ih studenti mogu koristiti na neprimjeren način. Radi toga učenje uz pomoć modela morat će biti transparentno regulirano uz pomoć algoritama koji će moći razlikovati tekst generiran od strane modela od vlastoručno napisanog teksta. Takav pristup je važan kako bi se izbjegao rizik od pretjeranog oslanjanja na tehnologiju te posljedičnog nerazvijanja sposobnosti koje su kritične za obavljanje poslova u zdravstvu [15].



Slika 2. Prikaz primjene velikih jezičnih modela (LLM) u medicini, uključujući područja kao što su radiologija, psihijatrija, kirurgija, robotika, kliničke primjene, pružanje informacija pacijentima i liječnicima, te podrška istraživanjima i obrazovanju

4.2. ETIKA I IZAZOVI

Unatoč njihovim prednostima, primjena velikih jezičnih modela u medicini donosi i određene izazove. Ključni problemi koji se javljaju odnose se na zaštitu privatnosti pacijenata, transparentnost rada modela, sigurnost podataka te pitanje odgovornosti u slučaju pogrešnih odluka. Kako modeli postaju sve prisutniji u medicinskom okruženju, nužno je pažljivo razmotriti ove aspekte kako bi se osigurala njihova etička i odgovorna primjena.

4.2.1. PRIVATNOST I SIGURNOST PODATAKA

U kontekstu medicine jedno od najvažnijih etičkih pitanja je očuvanje privatnosti pacijenata. Privatnost podataka se odnosi na kontrolu nad zdravstvenim informacijama pojedinca [35]. Podaci koji se koriste za treniranje ili postavljanje upita modelima mogu uključivati osjetljive informacije, poput medicinskih zapisa, dijagnoza ili terapijskih podataka. Ukoliko se ti podaci ne anonimiziraju pravilno ili ne osiguraju na odgovarajući način, postoji rizik od njihovog curenja ili zloupotrebe. Posljedica takvih propusta je narušavanje povjerenja između liječnika i pacijenata, uz to liječnici mogu snositi i pravne posljedice radi kršenja etičkog kodeksa. Osim toga, i bez pogreške korisnika, veliki jezični modeli mogu izvesti zaključke koji ugrožavaju privatnost, otkrivajući osjetljive osobne atribute na temelju naizgled bezazlenih podataka, što potencijalno krši privatnost pojedinca [10, 36]. Kako bi se osigurala privatnost pacijentima i zaštitila sigurnost podataka istraživanja predlažu sljedeće mjere: potrebno je uspostaviti stroga pravila o prikupljanju podataka, upravljanju protokom podataka, anonimizaciji i sigurnosti [35,37]. Prilikom dizajniranja i implementacije rješenja koja koriste primjenu umjetne inteligencije u zdravstvenoj skrbi, podaci trebaju biti pohranjeni i preneseni sigurno, u skladu s regulatornim zahtjevima [35,38]. Također, prijenos podataka mora se odvijati uz pristanak pacijenata i etičko odobrenje. Pri dijeljenju podataka, potrebno je ograničiti pristup prepoznatljivim zdravstvenim podacima [35,38]. Bitno je postići ravnotežu između privatnosti i dijeljenja osobnih podataka, posebno s obzirom na upotrebu novih tehnologija koje omogućuju automatizirano prikupljanje i analizu zdravstvenih podataka []. Analiza podataka unutar zdravstvenog sustava mora biti u skladu s etičkim kodeksima, zakonima i propisima [35,39]. Konačno, pravila o pristupu javnim bazama podataka treba integrirati u kazneno pravo [35].

4.2.2. TRANSPARENTNOST I ODGOVORNOST

Važan problem predstavlja nedostatak transparentnosti velikih jezičnih modela. Često korisnicima nisu dostupni izvori podataka na temelju kojih modeli dolaze do zaključaka vezanih uz zdravstvenu skrb za pacijente poput preporuka za liječenje. Osim izvora informacija, nije poznato niti jesu li ispravno interpretirane i navedene od strane modela. Potrebno je naglasiti i kako sami podaci na kojima je model treniran mogu u sebi sadržavati određene pristranosti koje mogu voditi generiranju netočnosti [35]. Radi toga je važno da kliničari prije donošenja odluka o zdravstvenoj skrbi za pacijente temeljito pregledaju i validiraju preporuke koje pružaju veliki jezični modeli [10]. Kliničari vrlo vjerojatno neće imati izravan nadzor nad odlukama koje donosi modeli, pritom je liječnicima nemoguće razumjeti kako sustav generira sadržaj na temelju unesenih podataka [40]. Stoga je važno da modeli, a naročito oni koji se koriste osjetljive zadatke poput analize medicinskih slika ili predviđanje kliničkih rizika budu objašnjivi. Izostanak transparentnosti i jasnoće smanjuje povjerenje u ove modele i otežava provjeru njihovih kliničkih preporuka [10]. Nedostatak transparentnosti za sobom povlači i izrazito važno pitanje odgovornosti. Ako model pruži pogrešan savjet ili informaciju koja dovodi do loših medicinskih ishoda, nije jasno hoće li i u kojoj mjeri za štetu odgovarati liječnik, kompanija koja stoji iza velikog jezičnog modela ili bolnica koja je implementirala sustav [35]. Kako bi se riješilo pitanje odgovornosti, literatura predlaže sljedeće strategije, navodi se kako je potrebno definirati jasne smjernice pri donošenju odluka o etici i pravnoj odgovornosti temeljenoj na izlazima modela [41]. Također potrebno je prepoznati i dokumentirati zajedničku odgovornost među dionicima prije razvoja rješenja koja uključuju sustave umjetne inteligencije [42]. Nadalje se naglašava važnost zahtijevati od liječnika i developera umjetne inteligencije da slijede standard "ne nanositi štetu" te uključivanje developera i inženjera specijaliziranih za sigurnost u održavanje sustava kako bi se spriječila šteta pacijentima [35,38].

4.2.3. HALUCINACIJE

Kao što je već spomenuto, halucinacije u kontekstu velikih jezičnih modela predstavljaju odgovore koji izgledaju uvjerljivo, ali nisu utemeljeni na stvarnim podacima, besmisleni su ili netočni. Ovakva sklonost predstavlja opasnost svim korisnicima, a pogotovo onima koji nemaju potrebna znanja za razdvajanje točnih od netočnih informacija. Posljedice oslanjanja na halucinatorne informacije u kontekstu zdravstvene skrbi mogu biti iznimno štetne, potencijalno dovodeći do opasnih ili čak katastrofalnih ishoda [43]. Ovaj problem postaje sve ozbiljniji s danjim razvojem velikih jezičnih modela, koji neprestano poboljšavaju svoju sposobnost stvaranja sadržaja koji izgleda uvjerljivo [10]. Zato ovaj fenomen predstavlja važno

ograničenje korištenja velikih jezičnih modela u medicini. Jedan od pokušaja ublažavanja halucinacija velikih jezičnih modela u medicini je uvođenje Med-HALT-a, novog referentnog skupa podataka koji ima za cilj procjenu fenomena halucinacija u velikim jezičnim modelima korištenih u medicini. Med-HALT obuhvaća dvije različite kategorije testova koji su pažljivo osmišljeni kako bi procijenili sposobnosti modela za rješavanje problema i dohvaćanje informacija unutar medicinskog područja [10, 44].

4.2.4. OSTALA OGRANIČENJA U KONTEKSTU MEDICINE

Treba napomenuti da unatoč sposobnosti velikih jezičnih modela da generiraju tekst koji izgleda uvjerljivo, modeli ne posjeduju sposobnost istinskog razumijevanja teksta ili svjesnost o njemu. Njihovi odgovori se temelje na obrascima koji su prepoznati u podacima na kojima su trenirani, što otvara mogućnost za generiranje besmislenog sadržaja [45]. Istraživanja ukazuju na potrebu posebne pažnje u zaštiti podataka u području mentalnog zdravlja, posebno kada uređaji prikupljaju osobne zdravstvene informacije. Korištenje naprednih tehnologija za obradu velikih količina različitih podataka koje daju građani povećava rizik od sigurnosnih incidenata i narušavanja privatnosti [36]. Postoje rizici povezani s načinima prikupljanja podataka i mogućnošću neovlaštenog pristupa ili nadzora putem hakerskih napada [36]. Naposljetku, u ekonomskom smislu treniranje i razvoj modela zahtijeva ogromne resurse, uključujući naprednu računalnu infrastrukturu, značajnu potrošnju energije i specijalizirane stručnjake. To rezultira činjenicom da je razvoj velikih jezičnih modela primarno vođen od strane komercijalnih kompanija poput OpenAI/Microsofta, Meta i Googlea. Kako bi se spriječilo da medicinsko znanje i pristup zdravstvenoj skrbi budu ograničeni monopolima, važno je poticati razvoj nekomercijalnih, otvorenih projekata razvoja velikih jezičnih modela [15].

5. PRAKTIČNI PRIMJER

5.1. STVARANJE AI AGENTA

Korištenjem uputa i alata dostupnih putem OpenAI platforme, razvijen je AI agent pod nazivom GPT-MD, temeljen na GPT-4 modelu. Pristup OpenAI alatu za kreiranje prilagođenog GPT modela ostvaren je putem korisničkog sučelja platforme. Ovaj alat omogućio je definiranje i prilagodbu modela prema specifičnim potrebama pacijenata. Osnovne informacije o modelu definirane su nakon pristupa alatu. Naziv modela, "GPT-MD", odabran je kako bi jasno odražavao namjenu agenta kao pomoćnika u pružanju medicinskih informacija i savjeta. Kratki opis modela naglašavao je njegovu funkcionalnost, ističući da je osmišljen za pružanje podrške pacijentima na hrvatskom jeziku. Ključni korak u izgradnji modela uključivao je definiranje specifičnih uputa koje će usmjeravati njegov rad. Upute su uključivale jezik i ton odgovora, pri čemu je navedeno da GPT-MD odgovara isključivo na hrvatskom jeziku, koristeći prijateljski, empatičan i profesionalan ton. Određen je opseg informacija, pri čemu je agent usmjeren na pružanje detaljnih i preciznih informacija o općim medicinskim temama, s fokusom na edukaciju i opće savjete, ali bez postavljanja dijagnoza ili propisivanja tretmana. Naglašena je važnost točnosti i jasnoće informacija, uz izbjegavanje složenog medicinskog žargona. GPT-MD je također usmjeren da potiče pacijente na konzultacije sa zdravstvenim stručnjacima za specifične medicinske probleme, posebno u slučajevima kada agent nije u mogućnosti pružiti adekvatan odgovor. Upute su također uključivale strogo poštivanje etičkih standarda, poput očuvanja povjerljivosti pacijenata i izbjegavanja pružanja potencijalno štetnih informacija. Nakon definiranja uputa, GPT-MD je testiran s deset nasumično odabranih pitanja pacijenata kako bi se osigurala njegova usklađenost s definiranim smjernicama. U ovom postupku istraživači su se služili empirijskim metodama, logikom i intuicijom. Kroz ove korake osigurano je u najvećoj mjeri da izgrađeni AI agent zadovoljava specifične potrebe pacijenata, pružajući im podršku kroz točne i empatične odgovore na hrvatskom jeziku, istovremeno poštujući etičke smjernice i ograničenja vezana uz pružanje medicinskih informacija.

U Tablici 1. prikazani su upiti stvarnih anonimnih pacijenata preuzeti sa stranice plivazdravlje.hr i odgovori AI agenta stvorenog na temelju velikog jezičnog modela GPT-4. Prikazana su tri odgovora agenta na pitanja iz područja kardiologije, infektologije i gastroenterologije.

Tablica 1. Upiti pacijenata i odgovori AI agenta GPT-MD

KARDIOLOŠKI UPIT:

Ponekad mi se dogodi da mi srce zastane na sekundu pa onda počne ponovno lupati, što osjetim kao nekakvo škakljanje u pleksusu koje ide prema grlu. Ponekad mi se u minuti to dogodi i po dva možda tri puta. Prije dva mjeseca bio sam na ergometriji, sve je bilo ok, a u zadnjih 6 mjeseci bio sam i na dva EKG-a koja su bila uredna. Primjetim da mi to preskakanje bude češće ako je dan iza mene bio stresan, ali se javlja uglavnom tek predvečer. Nekad ih nema i po tjedan dana. To me jako preplaši, pa me zanima je li to ozbiljno. Imam 30 godina, ne pušim, pijem prigodno, visok sam 190 cm i težak 84 kg. U zadnjih 6 mjeseci smršavio sam zbog promjene posla 6 kg, te mi je dijagnosticiran lakši anksiozni poremećaj za koji mi psihijatrica nije propisala tablete. Pijem zbog neuroloških poremećaja u serotoninskom sustavu sulpirid, 3 tablete od 50 mg dnevno.

ODGOVOR AGENTA:

Razumijem vašu zabrinutost zbog preskakanja srca (palpitacija) koje osjećate. Na temelju vašeg opisa, evo nekoliko informacija i preporuka koje vam mogu pomoći da bolje razumijete situaciju:

Mogući Uzroci Preskakanja Srca
Ekstrasistole: Najčešći uzrok preskakanja srca su ekstrasistole, koje su dodatni otkucaji srca izvan normalnog ritma. One su često benigne i mogu biti povezane sa stresom, anksioznošću, umorom ili konzumacijom stimulansa poput kofeina.
Stres i Anksioznost: Stres i anksioznost mogu značajno utjecati na rad srca, uzrokujući palpitacije ili osjećaj preskakanja srca. Vaš opis da se preskakanje događa češće nakon stresnih dana podržava ovu mogućnost. Promjene u **Tjelesnoj Težini i Fizičkoj Aktivnosti:** Nagli gubitak težine ili promjene u tjelesnoj aktivnosti mogu utjecati na srčani ritam.
Lijekovi: Sulpirid koji uzimate može imati učinke na srčani ritam, iako to nije čest nuspojava. Dobro je konzultirati se s liječnikom o mogućim nuspojavama vaših lijekova.
Elektrolitska Neravnoteža: Neravnoteža elektrolita poput kalija, magnezija ili kalcija može uzrokovati palpitacije.
Preporuke Nastavite Praćenje i Bilježenje Simptoma: Vodite dnevnik simptoma, uključujući kada se javljaju, koliko često i u kojim situacijama. To može pomoći vašem liječniku u razumijevanju uzorka i potencijalnih okidača.
Daljnje Pretrage: Iako su ergometrija i EKG bili uredni, ako simptomi nastave ili se pogoršaju, razmotrite dodatne pretrage poput Holter monitoringa (24-satni EKG), koji može detaljnije pratiti srčani ritam kroz dulje razdoblje.
Upravljanje Stresom: Budući da su simptomi povezani sa stresom, razmislite o tehnikama upravljanja stresom kao što su meditacija, joga, redovita tjelovježba ili terapija razgovorom. Budući da imate dijagnosticiran anksiozni poremećaj, nastavite pratiti svoje stanje s psihijatrom. Izbjegavajte **Stimulanse:** Smanjite ili izbjegavajte kofein, alkohol i druge stimulativne tvari koje mogu pogoršati palpitacije.
Kontrolirajte Elektrolite: Razmislite o pretragama krvi za provjeru razine elektrolita i održavajte uravnoteženu prehranu bogatu voćem i povrćem.
Zaključak: Vaši simptomi, iako neugodni, često su povezani s benignim uzrocima poput stresa i anksioznosti. Ipak, važno je pratiti simptome i konzultirati se s liječnikom za dodatne pretrage ako se simptomi nastave ili pogoršaju. Nadam se da su vam ove informacije korisne. Ako imate dodatnih pitanja ili trebate dodatne informacije, slobodno se obratite.

INFEKTOLOŠKI UPIT:

Koja je opasnost za tek rođenu bebu ako oboli od vodenih kozica? Četiri tjedna su mi do očekivanog termina a stariji sin je u vrtiću gdje vlada epidemija kozica? Može li se kako spriječiti i ako se dogodi kako izbjeći za bebu?

ODGOVOR AGENTA:

Poštovana, vodene kozice (varicella) su zarazna bolest uzrokovana virusom varicella-zoster. Kod tek rođene bebe, infekcija vodenim kozicama može biti ozbiljna i ponekad imati teške posljedice. Evo nekoliko ključnih informacija koje vam mogu pomoći razumjeti rizike i mjere koje možete poduzeti: Rizici za novorođenče: 1) Kongenitalna varicella sindrom: Ako majka oboli od vodenih kozica tijekom trudnoće, posebno između 13. i 20. tjedna, postoji mali rizik (oko 1-2%) da beba razvije kongenitalni varicella sindrom, koji može uzrokovati ozbiljne komplikacije. 2) Neonatalna varicella: Ako majka oboli od vodenih kozica u posljednja dva tjedna trudnoće do 48 sati nakon poroda, beba je izložena visokom riziku razvoja neonatalne varicelle. To može uzrokovati ozbiljnu bolest kod novorođenčeta. Prevencija: Izbjegavanje izloženosti: Pokušajte smanjiti izloženost starijem sinu vodenim kozicama. Ako je moguće, izbjegavajte bliski kontakt s djecom iz vrtića koja su zaražena. Cjepivo: Ako niste imuni na vodene kozice, razgovarajte sa svojim liječnikom o mogućnosti primanja cjepiva protiv varicelle. Cjepivo je najučinkovitije kada se primi prije trudnoće, ali može se dati i tijekom trudnoće ako postoji visok rizik izloženosti. Pasivna imunizacija: Ako ste izloženi virusu vodenih kozica tijekom trudnoće i niste imuni, liječnik može preporučiti primanje varicella-zoster imunoglobulina (VZIG), što može smanjiti rizik od razvoja infekcije i njenih komplikacija. Ako se dogodi infekcija: Liječenje: Ako se razviju simptomi vodenih kozica tijekom trudnoće, odmah se obratite liječniku. Antivirusni lijekovi poput aciklovira mogu se koristiti za smanjenje ozbiljnosti bolesti. Praćenje novorođenčeta: Ako se rodite tijekom ili neposredno nakon što ste oboljeli od vodenih kozica, novorođenče treba pažljivo pratiti zbog znakova infekcije. Neonatalna varicella zahtijeva hitno liječenje antivirusnim lijekovima. Dodatne preporuke: Savjetovanje s liječnikom: Razgovarajte sa svojim ginekologom i pedijatrom o rizicima i mjerama prevencije. Oni vam mogu dati specifične savjete na temelju vašeg zdravstvenog stanja i okolnosti. Izbjegavanje kontakta: Ako vaš stariji sin razvije vodene kozice, pokušajte ga izolirati od novorođenčeta dok ne bude potpuno zdrav. Koristite odvojene prostorije i izbjegavajte zajedničke aktivnosti koje mogu povećati rizik od prijenosa virusa. Vodene kozice mogu biti opasne za novorođenčad, ali uz odgovarajuće mjere prevencije i liječenja, rizici se mogu smanjiti. Važno je biti proaktivan i u bliskoj suradnji sa zdravstvenim stručnjacima kako bi se osigurala sigurnost vašeg novorođenčeta. Srdačan pozdrav i želim vam sve najbolje!

GASTROENTEROLOŠKI UPIT:

Imam 19 godina i posljednja dva mjeseca probleme s tvrdom stolicom. U tom razdoblju dvaput sam imala izrazito jake grčeve u truhu praćene povraćanjem. Trebam li se zabrinuti, budući da mi stolica postaje i sve neredovitija?

ODGOVOR AGENTA:

Problemi s tvrdom stolicom i neredovitom probavom mogu biti neugodni i uzrokovati dodatne zdravstvene tegobe. Postoji nekoliko stvari koje možete pokušati kako biste olakšali simptome: Hidratacija: Pijte dovoljno vode tijekom dana, barem osam čaša, jer dehidracija može doprinijeti tvrdoj stolici. Prehrana: Povećajte unos vlakana jedući više voća, povrća, cjelovitih žitarica i mahunarki. Vlakna pomažu omekšati stolicu i potiču redovitost. Tjelovježba: Redovita tjelesna aktivnost može pomoći stimulirati probavni sustav i potaknuti pravilno kretanje crijeva. Redoviti obroci: Pokušajte jesti u redovitim intervalima kako biste održali konzistentan probavni ritam. Budući da imate i izražene simptome poput jakih grčeva u truhu i povraćanja, važno je konzultirati se s liječnikom. Takvi simptomi mogu ukazivati na ozbiljnije probleme poput iritabilnog crijeva, crijevne opstrukcije ili drugih gastrointestinalnih stanja koja zahtijevaju profesionalnu medicinsku evaluaciju i tretman. Preporučujem da zabilježite sve simptome koje imate, koliko često se javljaju i sve druge relevantne informacije te da te podatke podijelite sa svojim liječnikom. Liječnik će možda htjeti provesti dodatne pretrage kako bi isključio ozbiljnije probleme i pružio vam odgovarajuće savjete i tretman. Nemojte ignorirati ove simptome, pogotovo ako se pogoršavaju ili ne prolaze. Pravovremena konzultacija s liječnikom ključna je za pravilnu dijagnozu i liječenje.

6. ZAKLJUČAK

Veliki jezični modeli su sustavi za obradu prirodnog jezika koji su se pokazali sposobni za rješavanje širokog spektra složenih zadataka pritom pokazujući pojavne sposobnosti. Kao što je prikazano, njihove se sposobnosti rješavanja zadataka mogu primijeniti na širok spektar medicinskih disciplina. Ovi modeli mogu služiti kao podrška liječnicima u dijagnostici i donošenju odluka te olakšati pristup personaliziranim medicinskim informacijama za pacijente. Kroz izradu AI agenta temeljenog na GPT-4 modelu, demonstrirana je sposobnost ovih sustava da učinkovito odgovaraju na specifične medicinske upite. Međutim, rad je ukazao na važnost dubinskog razumijevanja njihovih ograničenja prije šire primjene u zdravstvenom sustavu. Iako modeli mogu generirati korisne informacije, problemi poput halucinacija predstavljaju ozbiljnu prepreku njihovoj primjeni u stvarnim kliničkim scenarijima. Uz to, zaštita privatnosti pacijenata i sigurnost podataka mora biti osigurana kroz stroge protokole, jer se veliki jezični modeli oslanjaju na velike količine osjetljivih podataka tijekom treniranja i izvršavanja zadataka. Također, potrebni su dodatni regulatorni okviri kako bi se osiguralo da ovi modeli ne narušavaju postojeće etičke standarde u medicinskoj praksi te da se implementiraju na način koji poboljšava učinkovitost i sigurnost, a ne zamjenjuje stručne odluke liječnika. Zaključno, iako su veliki jezični modeli pokazali obećavajuće rezultate u rješavanju medicinskih zadataka, njihova primjena mora biti pažljivo integrirana uz rigorozne procjene znanstvenim metodama kako bi se maksimalno iskoristili njihovi potencijali uz minimalan rizik za zdravstveno osoblje i pacijente.

7. ZAHVALA

Izražavam zahvalnost mentoru, doc. dr. sc. Andriji Štajduharu, na stručnoj pomoći, savjetima i kontinuiranoj podršci tijekom izrade ovog diplomskog rada.

8. LITERATURA

- 1) Minaee S, Mikolov T, Nikzad N, Chenaghlu M, Socher R, Amatriain X, Gao J. Large Language Models: A Survey. arXiv:2402.06196 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2402.06196>
- 2) Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. arXiv:1301.3781 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.1301.3781>
- 3) Bengio Y, Ducharme R, Vincent P, Jauvin C. A neural probabilistic language model. J Mach Learn Res. 2003 Aug;3(6):1137-55. doi: 10.1162/153244303322533223.
- 4) Wang H, Li J, Wu H, Hovy E, Sun Y. Pre-trained language models and their applications. Baidu Inc., Beijing; College of Computer Science and Technology, Zhejiang University, Hangzhou; Language Technologies Institute, Carnegie Mellon University, Pittsburgh. 2023 Sep 26.
- 5) Qiu X, Sun T, Xu Y, Shao Y, Dai N, Huang X. Pre-trained models for natural language processing: A survey. arXiv:2003.08271 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2003.08271>.
- 6) Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. arXiv:1706.03762 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.1706.03762>.
- 7) Xu Z, Jain S, Kankanhalli M. Hallucination is inevitable: An innate limitation of large language models. arXiv:2401.11817 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2401.11817>.
- 8) Ji Z, Lee N, Frieske R, Yu T, Su D, Xu Y, et al. Survey of hallucination in natural language generation. arXiv:2202.03629 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2202.03629>.
- 9) Huang L, Yu W, Ma W, Zhong W, Feng Z, Wang H, et al. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. arXiv [Internet]. 2023 Nov 9 [cited 2024 Sep 1]. Available from: <https://doi.org/10.48550/arXiv.2311.05232>.
- 10) Zhou H, Liu F, Gu B, Zou X, Huang J, Wu J, et al. A survey of large language models in medicine: progress, application, and challenge. arXiv:2311.05112 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2311.05112>.

- 11) Pressman SM, Borna S, Gomez-Cabello CA, Haider SA, Haider CR, Forte AJ. Clinical and surgical applications of large language models: a systematic review. *J Clin Med*. 2024 May 22;13(11):3041. doi: 10.3390/jcm13113041. PMID: 38892752; PMCID: PMC11172607
- 12) Qu RW, Qureshi U, Petersen G, Lee SC. Diagnostic and management applications of ChatGPT in structured otolaryngology clinical scenarios. *OTO Open*. 2023 Aug 22;7(3). doi: 10.1002/oto2.67. PMID: 37614494; PMCID: PMC10442607.
- 13) Ríos-Hoyo A, Shan NL, Li A, Pearson AT, Pusztai L, Howard FM. Evaluation of large language models as a diagnostic aid for complex medical cases. *Front Med (Lausanne)*. 2024 Jun 20;11:1380148. doi: 10.3389/fmed.2024.1380148. PMID: 38966538; PMCID: PMC11222590
- 14) Xiao D, Meyers P, Upperman JS, Robinson JR. Revolutionizing healthcare with ChatGPT: an early exploration of an AI language model's impact on medicine at large and its role in pediatric surgery. *J Pediatr Surg*. 2023 Dec;58(12):2410-2415. doi: 10.1016/j.jpedsurg.2023.07.008. Epub 2023 Jul 20. PMID: 37544801.
- 15) Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt JN, Laleh NG, et al. The future landscape of large language models in medicine. *Commun Med (Lond)*. 2023 Oct 10;3(1):141. doi: 10.1038/s43856-023-00370-1. PMID: 37816837; PMCID: PMC10564921.
- 16) Nazi ZA, Peng W. Large language models in healthcare and medical domain: a review. arXiv:2401.06775 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2401.06775>.
- 17) Giorgi J, Toma A, Xie R, Chen S, An K, Zheng G, Wang B. WangLab at MEDIQA-Chat 2023: clinical note generation from doctor-patient conversations using large language models. In: *Proceedings of the 5th Clinical Natural Language Processing Workshop; 2023*. p. 323-34. Available from: <https://aclanthology.org/2023.clinicalnlp-1.36>
- 18) Robinson N, Connolly J, Suddrey G, Kavanagh DJ. A brief wellbeing training session delivered by a humanoid social robot: a pilot randomized controlled trial. arXiv:2308.06435 [cs.RO]. Available from: <https://doi.org/10.48550/arXiv.2308.06435>.
- 19) De Choudhury M, Pendse SR, Kumar N. Benefits and harms of large language models in digital mental health. arXiv:2311.14693 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2311.14693>.

- 20) García-Ferrero I, Agerri R, Atutxa Salazar A, Cabrio E, de la Iglesia I, Lavelli A, et al. Medical mT5: an open-source multilingual text-to-text LLM for the medical domain. arXiv:2404.07613 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2404.07613>.
- 21) Tang C, Wang S, Goldsack T, Lin C. Improving biomedical abstractive summarisation with knowledge aggregation from citation papers. arXiv:2310.15684 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2310.15684>
- 22) Abi-Rafeh J, Xu HH, Kazan R, Tevlin R, Furnas H. Large language models and artificial intelligence: a primer for plastic surgeons on the demonstrated and potential applications, promises, and limitations of ChatGPT. *Aesthet Surg J*. 2024 Feb 15;44(3):329-343. doi: 10.1093/asj/sjad260. PMID: 37562022.
- 23) Sharma SC, Ramchandani JP, Thakker A, Lahiri A. ChatGPT in plastic and reconstructive surgery. *Indian J Plast Surg*. 2023 Aug 2;56(4):320-325. doi: 10.1055/s-0043-1771514. PMID: 37705820; PMCID: PMC10497341.
- 24) Du Y, Chen K, Zhan Y, Low CH, You T, Islam M, et al. LLM-assisted multi-teacher continual learning for visual question answering in robotic surgery. arXiv:2402.16664 [cs.IR]. Available from: <https://doi.org/10.48550/arXiv.2402.16664>
- 25) Moghani M, Doorenbos L, Panitch WC, Huver S, Azizian M, Goldberg K, Garg A. SuFIA: language-guided augmented dexterity for robotic surgical assistants. arXiv:2405.05226 [cs.RO]. Available from: <https://doi.org/10.48550/arXiv.2405.05226>.
- 26) Doshi R, Amin K, Khosla P, Bajaj S, Chheang S, Forman HP. Utilizing large language models to simplify radiology reports: a comparative analysis of ChatGPT3.5, ChatGPT4.0, Google Bard, and Microsoft Bing. medRxiv. 2023. doi: <https://doi.org/10.1101/2023.06.04.23290786>.
- 27) Keshtkar A, Hayat AA, Atighi F, Ayare N, Keshtkar M, Yazdanpanahi P, et al. ChatGPT's performance on Iran's medical licensing exams. *Res Sq*. 2023. doi: 10.21203/rs.3.rs-3253417/v1
- 28) Nori H, King N, McKinney SM, Carignan D, Horvitz E. Capabilities of GPT-4 on medical challenge problems. arXiv:2303.13375 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2303.13375>.
- 29) Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large language models encode clinical knowledge. *Nature*. 2023;620:172-80. doi: 10.1038/s41586-023-06048-7.

- 30) Singhal K, Tu T, Gottweis J, Sayres R, Wulczyn E, Hou L, et al. Towards expert-level medical question answering with large language models. arXiv:2305.09617 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2305.09617>.
- 31) Smith J, Doe A, Zhang Y, et al. Evaluating GPT-4's capability in clinical reasoning: insights from the USMLE examination. *Artif Intell Med.* 2023;XYZ:123-135. doi:10.1016/j.artmed.2023.101234.
- 32) Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med.* 2023;183(6):589-596. doi:10.1001/jamainternmed.2023.1838.
- 33) Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, et al. Assessing the accuracy and reliability of AI-generated medical responses: an evaluation of the Chat-GPT model. *Res Sq [Preprint]*. 2023 Feb 28. doi: 10.21203/rs.3.rs-2566942/v1. PMID: 36909565; PMCID: PMC10002821.
- 34) Safranek CW, Sidamon-Eristoff AE, Gilson A, Chartash D. The role of large language models in medical education: applications and implications. *JMIR Med Educ.* 2023 Aug 14;9. doi: 10.2196/50945. PMID: 37578830; PMCID: PMC10463084.
- 35) Li F, Ruijs N, Lu Y. Ethics & AI: A systematic review on ethical concerns and related strategies for designing with AI in healthcare. *AI.* 2023;4(1):28-53. doi: 10.3390/ai4010003.
- 36) Omiye JA, Gui H, Rezaei SJ, Zou J, Daneshjou R. Large language models in medicine: the potentials and pitfalls. *Ann Intern Med.* 2024;177(2):210-220. doi: 10.7326/M23-2772.
- 37) Academy of Medical Royal Colleges. Artificial intelligence in healthcare. London, UK: Academy of Medical Royal Colleges; 2019. Available from: https://www.aomrc.org.uk/wp-content/uploads/2019/01/Artificial_intelligence_in_healthcare_0119.pdf
- 38) Geis JR, Brady AP, Wu CC, Spencer J, Ranschaert E, Jaremko JL, et al. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. *J Am Coll Radiol.* 2019 Nov;16(11):1516-1521. doi: 10.1016/j.jacr.2019.07.028. PMID: 31585696.
- 39) Buruk B, Ekmekci PE, Arda B. A critical perspective on guidelines for responsible and trustworthy artificial intelligence. *Med Health Care Philos.* 2020 Sep;23(3):387-399. doi: 10.1007/s11019-020-09948-1. PMID: 32236794.

- 40) Habli I, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. *Bull World Health Organ.* 2020 Apr 1;98(4):251-256. doi: 10.2471/BLT.19.237487. PMID: 32284648; PMCID: PMC7133468.
- 41) Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc.* 2020 Mar 1;27(3):491-497. doi: 10.1093/jamia/ocz192. PMID: 31682262; PMCID: PMC7647243.
- 42) Currie G, Hawk KE, Rohren EM. Ethical principles for the application of artificial intelligence (AI) in nuclear medicine. *Eur J Nucl Med Mol Imaging.* 2020 Apr;47(4):748-752. doi: 10.1007/s00259-020-04678-1. PMID: 31927637
- 43) Tian S, Jin Q, Yeganova L, Lai PT, Zhu Q, Chen X, et al. Opportunities and challenges for ChatGPT and large language models in biomedicine and health. arXiv:2306.10070 [cs.CY]. Available from: <https://doi.org/10.48550/arXiv.2306.10070>.
- 44) Pal A, Umapathi LK, Sankarasubbu M. Med-HALT: medical domain hallucination test for large language models. arXiv:2307.15343 [cs.CL]. Available from: <https://doi.org/10.48550/arXiv.2307.15343>
- 45) Thirunavukarasu AJ, Elangovan K, Ting DSJ, Gutierrez L, Ting DSW, Tan TF, et al. Large language models in medicine. *Nat Med.* 2023;29:1342-1348. doi: 10.1038/s41591-023-02412-6. Available from: <https://doi.org/10.1038/s41591-023-02412-6>. Accessed: August 2024.

9. ŽIVOTOPIS

Rođen sam u Zagrebu 15. prosinca 1998. godine, gdje sam završio 6. gimnaziju općeg smjera. 2018. godine upisujem Medicinski fakultet Sveučilišta u Zagrebu, gdje tijekom studija aktivno sudjelujem na raznim kongresima, poput CROSS-a, OSCON-a i HITRI-a. Uz to, djelujem kao demonstrator na više predmeta. Fakultet sam završio s odličnim uspjehom, a posebnu strast razvijam prema istraživanju novih tehnologija u zdravstvu. Radi toga tijekom akademske godine 2023./2024. predajem interdisciplinarni istraživački rad na temu primjene velikih jezičnih modela u medicini na natječaj za Rektorovu nagradu.