

LLM + DEX: BRŽA I OBJAŠNJIVA KONSTRUKCIJA HIJERARHIJSKIH MODELA ODLUČIVANJA

LLM + DEX: FASTER AND EXPLAINABLE CONSTRUCTION OF HIERARCHICAL DECISION-MAKING MODELS

SANDRO RADOVANOVIĆ¹, BORIS DELIBAŠIĆ², MARKO BOHANEĆ³, MILIJA SUKNOVIĆ⁴

¹ Univerzitet u Beogradu, Fakultet organizacionih nauka, Beograd, sandro.radovanovic@fon.bg.ac.rs,  0000-0002-9975-8844

² Univerzitet u Beogradu, Fakultet organizacionih nauka, Beograd, boris.delibasic@fon.bg.ac.rs,  0000-0002-6153-5119

³ Institut Jožef Stefan, Ljubljana, Beograd, marko.bohanec@ijs.si,  0000-0003-4317-2833

⁴ Univerzitet u Beogradu, Fakultet organizacionih nauka, Beograd, milija.suknovic@fon.bg.ac.rs,  0009-0003-0449-5258

Rezime: Ovaj rad istražuje mogućnosti integracije velikih jezičkih modela u DEX metodologiju za višekriterijumsko odlučivanje. DEX modeli su interpretabilni, hijerarhijski organizovani sistemi odlučivanja zasnovani na kvalitativnim pravilima i preferencijalnoj logici, ali njihova izgradnja zahteva vreme i ekspertsko znanje. Sa druge strane, LLM modeli, poput ChatGPT, poseduju sposobnost razumevanja i generisanja strukture odlučivanja na osnovu opisa problema i podataka. U ovom radu predlažemo metod koji koristi LLM za izbor relevantnih atributa, konstrukciju hijerarhije nadkriterijuma i generisanje semantičkih objašnjenja, dok se pravila odlučivanja uče iz podataka uz poštovanje uslova monotonosti. Na eksperimentalnom primeru modelovanja rizika skijaških povreda, pokazujemo da ovako dobijeni modeli postižu performanse uporedive sa DDEX modelom i klasičnim modelima mašinskog učenja, uz značajno bržu izradu i veću interpretabilnost. Predloženi pristup otvara mogućnosti za dalju automatizaciju i razvoj inteligentnih agenata za donošenje odluka.

Ključne reči: Veliki jezički modeli, DEX, Interpretabilni modeli odlučivanja, Sistemi za podršku odlučivanju.

Abstract: This paper explores the possibilities of integrating Large Language Models into the DEX methodology for multi-criteria decision-making. DEX models are interpretable, hierarchically organized decision systems based on qualitative rules and preferential logic, but their construction requires time and domain expertise. On the other hand, LLMs, such as ChatGPT, possess the ability to process text and generate decision structures based on problem descriptions and data. In this paper, we propose a method that uses LLMs to select relevant attributes, construct a hierarchy of aggregate criteria, and generate semantic explanations, while decision rules are learned from data with monotonicity constraints. On an experimental case of ski injury risk modeling, we demonstrate that the resulting models achieve performance comparable to the DDEX model and standard machine learning models, with significantly faster development and improved interpretability. The proposed approach opens opportunities for further automation and the development of intelligent agents for decision-making support.

Keywords: Large language models, DEX, Interpretable decision models, Decision support systems.

1. UVOD

Sistemi za podršku odlučivanju (eng. *Decision Support Systems*, ili skraćeno *DSS*) predstavljaju računarske alate namenjene pomaganju ljudima u donošenju odluka u složenim i neizvesnim uslovima (Burststein et al., 2008). Za razliku od automatizovanih sistema odlučivanja koji samostalno generišu rešenja, *DSS* sistemi imaju ulogu podrške ili asistenta, tj. oni korisnicima pružaju strukturisane informacije, analitičku podršku i logičke okvire, ali krajnja odluka ostaje u rukama čoveka. Uloga ovih sistema posebno dolazi do izražaja kada je potrebno doneti odluku u domenu gde ne postoje jednostavne metrike uspeha, gde su vrednosti subjektivne, a ciljevi međusobno konfliktni (Turban et al., 2015). Primena *DSS* sistema obuhvata širok spektar oblasti, od strateškog menadžmenta i javne uprave (Delibašić, Radovanović, and Vukanović, 2023), do agronomije, zdravstva i obrazovanja. U svim tim slučajevima, zajednički imenilac je potreba da se prikupe, obrade i povežu različiti izvori podataka, ekspertsko znanja i preferencije donosioca odluke (Turban et al., 2015). Potom, umesto da zameni ljudsko razmišljanje, *DSS* sistemi bi trebalo da su dizajnirani da ga ojačaju, pritom obezbeđujući transparentnost, doslednost i sistematičnost u procesu donošenja

odluka.

Takođe, u poređenju sa pristupima baziranim na mašinskom učenju, koji su usmereni na otkrivanje zakonitosti u podacima i koji često funkcionišu kao "*crne kutije*", sistemi za podršku odlučivanju su usmereni na interpretabilnost, strukturiranost i logički zasnovano modelovanje znanja (Bohanec, 2021). Stoga, DSS sistemi kao što je DEX (Decision EXpert) oslanjaju se na eksplicitno definisane kriterijume, hijerarhije i pravila odlučivanja, omogućavajući korisnicima uvid u tok zaključivanja, kao i razloge iza svake preporuke. Na taj način, DSS podržava proces donošenja odluka ne samo kroz rezultate, već i kroz razumevanje i objašnjenje. U radu (Bohanec, 2021), istaknuto je da bi kvalitetan sistem za podršku odlučivanju trebalo da zadovolji takozvanih 5C uslova. To su: **tačnost** (eng. *Correctness*), odnosno da bi svaki DSS trebalo da pruža tačne i validne informacije relevantne za problem odlučivanja; **potpunosti** (eng. *Completeness*), tj. da su obuhvaćeni svi relevantni aspekti problema, kao i da sistem pruža odgovor za sve moguće ulaze; **doslednost** (eng. *Consistency*), misleći da postoji logička i preferencijalna usklađenost modela; **razumljivost** (eng. *Comprehensibility*) koja se odnosi na lakoću kojom korisnik može razumeti predstavljene informacije; i **prikladnost** (eng. *Convenience*) odnoseći se na dostupnost i pravovremenost informacija, prilagođenih konkretnom zadatku i korisniku.

Uprkos jasnim kriterijumima kvaliteta, izgradnja DSS sistema, tačnije DEX modela, koji zadovoljavaju svih 5C zahteva predstavlja značajan izazov. Potrebno je vreme, specifično znanje za domen problema i pažljivo modeliranje kriterijuma, pravila i hijerarhija odlučivanja. Ovaj proces često zahteva intenzivnu saradnju između eksperata i modelara, što ograničava skalabilnost i brzinu razvoja ovakvih sistema.

Sa druge strane veliki jezički modeli (eng. *Large Language Models*, skraćeno *LLM*), poput *ChatGPT*-ja ili *DeepSeek*-a, nude potencijalno rešenje za ove izazove. Zahvaljujući sposobnosti razumevanja prirodnog jezika, analiziranja problema i generisanja relevantnog sadržaja, *LLM* modeli mogu pomoći u formulisanju kriterijuma, strukturiranju modela i čak u simulaciji procesa odlučivanja. Na taj način, *LLM* može značajno smanjiti trošak i kompleksnost izrade DSS sistema, ubrzati njihov razvoj i učiniti ih dostupnijim široj grupi korisnika. Ovaj rad predstavlja početni, pozicioni korak ka istraživanju mogućnosti integracije velikih jezičkih modela u proces izgradnje DEX sistema za podršku odlučivanju. Cilj rada je da identifikuje potencijalne tačke povezivanja između ove dve paradigme i otvori prostor za dalja istraživanja u pravcu automatizacije i proširenja DSS metodologije kroz savremene alate veštačke inteligencije.

2. TEORIJSKE OSNOVE

U ovom delu opisani su ukratko DEX metodologija i veliki jezički modeli, te je diskutovana njihova komplementarnost.

2.1. Metodologija DEX

Jedan od najpoznatijih i najprimenjivanih pristupa u okviru sistema za podršku odlučivanju jeste DEX metodologija. DEX (eng. Decision EXpert) je metod višekriterijumske analize zasnovan na simboličkom, kvalitativnom modelovanju, koje omogućava izražavanje znanja specifičnih za domen kroz hijerarhijsku strukturu atributa i logički definisana pravila odlučivanja (Bohanec, 2022). Za razliku od kvantitativnih metoda, DEX koristi diskretne (kategoričke ili verbalne) vrednosti kriterijuma, čime omogućava donošenje odluka čak i u situacijama kada precizni numerički podaci nisu dostupni ili nisu poželjni. Iako se nekada korišćenje isključivo kvalitativnih skala vrednosti može smatrati nedostatkom, način na koji ljudi percipiraju vrednosti kriterijuma najčešće jesu u vidu kategorija vrednosti zbog nedostatka kognitivnih mogućnosti da posmatraju i poredi veliki broj različitih vrednosti (Moscato, 2021). Sa druge strane, moguće je i iz podataka, ako su dostupni, odrediti kategorije vrednosti i njihovu orijentaciju (Radovanović, Bohanec, and Delibašić, 2023).

DEX modeli se sastoje od atributa, tj. kriterijuma, koji su organizovani u vidu stablu odlučivanja, gde se složeni kriterijumi (nadkriterijumi) formiraju agregacijom jednostavnijih (ulaznih ili drugih složenih) kriterijuma, dok je u korenu stabla ciljni atribut, odnosno atribut u kojem se donosi odluka. Svaki složeni atribut ima pridruženu tabelu pravila odlučivanja (eng. decision table), koja definiše izlaznu vrednost na osnovu kombinacija ulaznih vrednosti njegovih podatributa. Na taj način, model funkcioniše kao sistem lokalno definisanih pravila koji se evaluira odozdo-nagore kroz hijerarhiju, sve do konačne odluke ili preporuke. Zapravo, tabela pravila odlučivanja ima ulogu računanja korisnosti alternative za zadati kriterijum ili ciljni atribut, te kao takav ima ulogu kompozitnog indikatora koji predstavlja očekivanu korisnost alternative (za zadati kriterijum ili ciljni atribut).

Ključna karakteristika DEX metodologije je uređenost vrednosti atributa, što znači da su vrednosti svakog atributa raspoređene u logički niz (npr. loš, srednji, dobar) koji odražava preferencije korisnika. Ova uređenost omogućava primenu pravila konzistentnosti i preferencijalne logike, čime se osigurava da model bude stabilan, predvidiv i u skladu sa očekivanjima korisnika (Bazerman and Moore, 2012). Upravo zbog ove osobine, DEX modeli su naročito pogodni za situacije u kojima su transparentnost i interpretabilnost odluke od presudne

važnosti. Bitno je napomenuti da korisnik može koristiti i neuređene skale vrednosti, ali u tom slučaju nije moguće primenjivati pravila konzistentnosti.

2.2. Veliki jezički modeli

Drugi aspekt koji analiziramo su veliki jezički modeli. Oni predstavljaju napredne modele veštačke inteligencije trenirane na ogromnim količinama tekstualnih podataka sa ciljem da razumeju, generišu i transformišu prirodni jezik. Kao takvi, zasnovani na arhitekturi transformera i sposobni su da proizvode koherentan i kontekstualno relevantan tekst, odgovaraju na pitanja, sažimaju sadržaj, formulišu predloge i vode dijalog na način sličan čoveku (Naveed et al., 2023). Ono što ih posebno izdvaja kao alate u poređenju sa drugim alatima veštačke inteligencije je njihova mogućnosti koja ne proističe iz eksplicitno programiranih pravila, već iz učenja zakonitosti i značenja iz velikog korpusa tekstova, te se neretko kaže da su u stanju da "prevare" čoveka da razgovara sa drugim čovekom, tj. da su položili su Turingov test (Jones and Bergen, 2025).

Pored toga što funkcionišu kao generativni modeli, veliki jezički modeli se sve više posmatraju i kao sistemi za upravljanje znanjem, jer u svojoj strukturi implicitno integrišu ogromnu količinu činjenica, pojmova i odnosa između njih (He et al., 2024). Iako nisu klasične baze znanja, LLM modeli mogu obavljati zadatke kao što su ekstrakcija informacija, povezivanje pojmova, preformulisanje i organizacija znanja, čime postaju funkcionalni alati za rad sa složenim informacijama. Zapravo, sistemi za upravljanje znanjem obuhvataju procese prikupljanja, organizovanja, skladištenja i distribuiranja znanja unutar organizacije ili ekspertskog sistema (Akerkar and Sajja, 2009). Kao takvi, sistemi za upravljanje znanjem se oslanjaju na strukturirane baze podataka, ontologije ili ekspertna pravila kako bi omogućili efikasno pronalaženje i primenu relevantnih informacija, i posledično donošenje odluka. Nasuprot tome, LLM-ovi predstavljaju drugačiji pristup modelovanja i upravljanja znanjem. LLM-ovi ne upravljaju znanjem eksplicitno, već ga modeluju implicitno kroz zakonitosti u podacima i semantičke relacije naučene iz ogromnog korpusa tekstualnih podataka. Uprkos tome što ne poseduju eksplicitno strukturiranu bazu znanja kao klasični sistemi, LLM-ovi se mogu funkcionalno posmatrati kao alati za upravljanje znanjem jer poseduju sve aktivnosti koje jedan sistem za upravljanje znanjem mora da poseduje.

Zahvaljujući sposobnosti da razumeju i generišu tekst, LLM modeli mogu izdvajati relevantne informacije iz neformalnih opisa, identifikovati kriterijume, definicije, čak i uzročno-posledične veze (Meng et al., 2022) što odgovara osobinama **ekstrakcije znanja, povezivanje pojmova i strukturiranja informacije**. Osim toga, LLM modeli mogu generisati i **objašnjavati znanje**, formulišući sažetke, alternativne formulacije i preporuke, čime olakšavaju prenos i deljenje znanja između učesnika u procesu odlučivanja. Ipak, pri korišćenju LLM-ova potrebno je biti veoma oprezan, jer su u stanju da generišu uverljive, ali netačne ili izmišljene informacije. To je fenomen poznat kao *halucinacija* ili *konfabulacija* modela (Barros, 2025).

U kontekstu sistema za podršku odlučivanju, navedene sposobnost LLM-a se mogu iskoristiti za generisanje kriterijuma, objašnjenja, pa čak i pravila odlučivanja, čime se LLM pozicionira kao alat za podršku donosiocima odluka pri konstrukciji i razumevanju modela odlučivanja. Međutim, ukoliko imamo dostupne podatke o problemu koji rešavamo, možemo pronaći kombinaciju u kojoj ćemo uzeti najbolje iz oba sveta. Velike jezičke modele da pomognu u konstruisanju hijerarhije kriterijuma, a podatke da popunimo tabelu pravila odlučivanja.

3. KOMBINOVANJE LLM-A I DEX METODOLOGIJE

U ovom radu želimo da predložimo modifikaciju *DIDEX* (Radovanović, Bohanec, and Delibašić, 2023) metodologije koja je kompletan postupak pravljenja *DEX* modela zasnivala na istorijskim podacima. U ovom istraživanju želimo da pomognemo postupak korišćenjem LLM modela. Tačnije, predloženi pristup se oslanja na integraciju LLM-ova i formalnog okvira DEX metodologije za generisanje hijerarhije modela odlučivanja zasnivajući se na neformalnom opisu problema, ali i dostupnim podacima. Model koristi sposobnost LLM-a da na osnovu opisa problema i atributa automatski predloži strukturisanu hijerarhiju kriterijuma, kao i semantički opravdana imena i objašnjenja za nove, složene attribute. Ovaj korak je srž *DEX* metodologije i vremenski najzahtevniji korak (Bohanec, 2022), kao i računski najkompleksniji u *DIDEX* metodologiji (Radovanović, Bohanec, and Delibašić, 2023).

Proces započinje tako što korisnik definiše opis problema, kao i opis podataka, gde svaki atribut ima naziv i kratak opis. LLM se tada koristi kao domenski ekspert koji, na osnovu datog opisa i značenja atributa, identifikuje koje attribute treba uključiti u model i kako ih međusobno povezati u strukturisanu formu. Odnosno, prvi korak je izbor informativnih kriterijuma. Formalno, putem strukturnog izlaza koji neki LLM-ovi podržavaju podskup, LLM (tj. virtuelni donosilac odluke) bira podskup od najviše k atributa koji su najznačajniji za konkretan problem odlučivanja. Ovaj postupak je veoma bitan jer *DEX* modeli bi trebalo da predstavljaju dovoljno kompleksne, ali ne previše kompleksne modele odlučivanja. Prosleđivanjem ogromnog broj kriterijuma rezultujući *DEX* model ne bi bio razumljiv donosiocima odluke zbog ili prevelike dubine hijerarhije ili zbog prevelikih tabela pravila odlučivanja koja će se neminovno

pojavit sa velikim brojem atributa.

Nakon izbora informativnih kriterijuma, u narednim iteracijama *LLM* predlaže **spajanje parova kriterijuma** u nove, složene kriterijume. Svaki predlog sadrži: (1) nazive dva kriterijuma za spajanje, (2) novo ime agregiranog kriterijuma i (3) tekstualno objašnjenje koje opisuje semantičku povezanost između izabranih kriterijuma i opravdanost njihove integracije. Na taj način, model gradi hijerarhiju odozdo nagore, sve dok se ne dobije jedinstven atribut na vrhu koji predstavlja konačnu odluku. Kako bismo obezbedili da *LLM* uvek vraća iste rezultate, korišćen je OpenAI GPT 4o model gde je temperatura spuštена na vrednost veoma bliskoj nuli (0,001). Nakon definisanja strukture *DEX* modela, svaki čvor u hijerarhiji predstavlja agregaciju dva kriterijuma,

za koje je potrebno **konstruisati tabelu pravila odlučivanja**. Pravila se uče direktno iz podataka, uz poštovanje uslova monotonosti koji su svojstveni *DEX* metodologiji, što znači da se povećanjem vrednosti ulaznih kriterijuma ne sme doći do degradacije izlazne vrednosti.

Za izbor najboljeg izlaza za svaku kombinaciju vrednosti atributa koristi se verovatnoća izlazne klase (eng. *Class Confidence Proportion*), koja je definisana kao:

$$p(c | a, b) = \frac{N_{abc}}{N_{ab}} \quad (1)$$

gde je N_{abc} broj instanci sa kombinacijom vrednosti a i b koje pripadaju klasi c , dok je N_{ab} ukupan broj instanci sa tom kombinacijom ulaza.

Da bi se postigla stabilnost i robusnost pravila, dodatno se koristi *donja granica pouzdanosti* (eng. *Lower Confidence Bound*, skraćeno *LCB*) za proporciju p , prema Vilsonovoj formuli (Rink and Brannath, 2025):

$$LBC(p) = \frac{p + \frac{z^2}{n} - z \sqrt{\frac{p(1-p)}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}} \quad (2)$$

gde je z kritična vrednost standardne normalne raspodele za zadati nivo poverenja (npr. $z \approx 1.96$ za 95%), a n je broj posmatranja za datu kombinaciju ulaza. Na taj način, za svaku kombinaciju atributa bira se ona klasa c za koju je *LCB* najviši, čime se favorizuju stabilne i konzervativne odluke u uslovima male frekvencije podataka. Nakon konstruisanja svake tabele pravila, a zarad očuvanja doslednosti u odlučivanju i preferencijalne konzistentnosti (što je ključno za *DEX* modele), vrši se provera i korekcija radi obezbeđivanja *monotonosti* pravila odlučivanja.

Trebalo bi napomenuti ovo nije prvi pristup u kombinovanju *LLM*-ova i *DEX* metodologije, već da postoji rad (Radovanović, Delibašić, and Vukanović, 2024) koji koristi *LLM* da konstruiše hijerarhiju u potpunosti, a tabele pravila odlučivanja se konstruišu iz podataka. Predložena metodologija u ovom radu se razlikuje zato što se hijerarhija konstruiše iz većeg broja uzastopnih upita *LLM*-a, te konstruisanje strukturnog izlaza i, na kraju, promenama u pravljenju tabele pravila odlučivanja koja obezbeđuju preferencijalnu konzistentnost, a koriste *LCB* kao metoda za određivanje vrednosti izlaza ciljnog atributa ili nadkriterijuma.

4. REZULTATI I DISKUSIJA

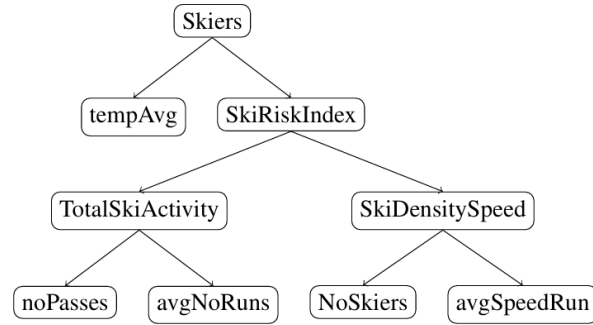
U cilju poređenja modela i rezultata, koristili smo identične podatke i identičnu postavku kao u (Radovanović, Bohanec, and Delibašić, 2023). Odnosno, koristili smo podatke o povreda skijaša na skijalištu Kopaonik u pet uzastopnih sezona. Podaci opisuju karakteristike skijališta, vremenske uslove i nivo aktivnosti korisnika i to su:

1. numSkiers – broj skijaša na skijalištu dnevno,
2. numPasses – broj transporta žičarom,
3. utilization – iskorišćenost kapaciteta žičara,
4. avgNoRuns – prosečan broj vožnji po skijašu,
5. avgAvgTimeRun, avgMinTimeRun, avgMaxTimeRun – različiti aspekti vremena koje skijaši provode u jednoj vožnji,
6. avgAvgElevation, avgMinElevation, avgMaxElevation – vertikalna dužina spusta,
7. avgSpeedRun – prosečna brzina skijanja,
8. tempAvg – prosečna dnevna temperatura,
9. windSpeed – brzina vetra,
10. cloudiness – oblačnost (na skali od 0 do 10),
11. numInjuries – broj prijavljenih povreda na dnevnom nivou podeljene u tri kategorije.

Takođe smo, kao i u originalnom radu, koristili diskretizaciju sa jednakim brojem instanci po kategoriji (eng. *equal size binning*) kako bismo dobili stabilnija pravila odlučivanja. Odnosno, podrška (eng. *support*) svakog pravila bi trebalo da je veća ovim pristupom u poređenju sa drugim metodama diskretizacije, čime se povećava pouzdanost dobijenih odluka.

Rezultati su evaluirani upotrebom desetostruke unakrsne validacije (eng. *cross validation*) gde su merene tačnost, preciznost, odziv i F1 mera. Prilikom izbora informativnih atributa, isprobane su verzije sa od 5 do 7 kriterijuma, a u radu ćemo prikazati samo model sa 5 kriterijuma (koja je prikazana na Slici 1:) kako bi bio

uporediv sa modelom dobijenim u (Radovanović, Bohanec, and Delibašić, 2023).



Slika 1: LLM dobijena DEX hijerarhija

U poređenju sa *DIDEX* modelom iz prethodnog rada, novi model zasnovan na *LLM*-u koristi sličan skup ulaznih atributa, ali uz određene razlike u njihovom izboru i organizaciji. Dok *LLM* model koristi attribute *tempAvg*, *NoSkiers*, *noPasses*, *avgNoRuns* i *avgSpeedRun*, *DIDEX* model koristi *NoSkiers*, *tempAvg*, *utilization*, *noPasses* i *windSpeed*. Odnosno, novi model modeluje vremenske uslove i rizik (kao kombinaciju skijaških aktivnosti i gustine) kao ključne faktore za rizik od skijaških povreda, dok *DIDEX* model vremenske uslove kombinuje sa brojem skijaša za modelovanje stanja skijališta, te njih sa brojem spusteva. Zapravo, skijaške povrede su veoma asocirane sa povećanom aktivnošću i gustinom skijaša na skijalištu. Naime, skijaške povrede se najčešće događaju u uslovima kada je skijalište prenatrpano i kada je rizik od sudara ili izbegavanja sudara veliki (Carus and Castillo, 2021), kada se formiraju manje grupe skijaša (Delibašić, Radovanović, Jovanović, et al., 2024) ili kada ima više neiskusnih skijaša.

Osim izbora atributa, ključna razlika ogleda se u načinu na koji su ovi atributi organizovani u hijerarhiji. *LLM* automatski konstruiše nadkriterijume *SkiActivityIndex* i *SkiDensityIndex* i daje njihova tumačenja. Ovi agregirani kriterijumi predstavljaju semantički utemeljene iz baze znanja *LLM*-a, čime donosilac odluke može da posmatra i diskutuje same izbore i time pomogne proces kreiranja *DEX* modela. Pored hijerarhijske strukture, koristeći podatke dobijamo tabele pravila odlučivanja, pri čemu su sva pravila monotona. Na kraju, u poređenju sa *DIDEX* metodologijom, ceo proces je realizovan drastično brže sa *LLM*-om.

Tabela 1: Evaluacija modela za skijaške povrede (*) Rezultati preuzeti iz (Radovanović, Bohanec, and Delibašić, 2023).

Model	Accuracy	Precision	Recall	F1
<i>LLM</i> – 5 kriterijuma	65.13 ± 6.34	63.51 ± 6.98	66.78 ± 5.74	63.06 ± 6.00
<i>LLM</i> – 6 kriterijuma	63.96 ± 5.97	63.11 ± 6.09	66.10 ± 4.62	63.13 ± 5.38
<i>LLM</i> – 7 kriterijuma	65.62 ± 4.43	64.54 ± 3.98	66.83 ± 4.75	64.86 ± 3.90
<i>DIDEX</i> *)	65.73 ± 12.16	64.51 ± 10.67	66.30 ± 7.37	64.07 ± 9.10
Stablo odlučivanja*)	63.96 ± 5.22	63.63 ± 5.50	65.72 ± 4.47	64.66 ± 5.09
Model nasumičnih stabala odlučivanja*)	66.27 ± 5.39	64.94 ± 5.94	67.32 ± 4.84	67.32 ± 4.34
Logistička regresija*)	63.32 ± 5.14	62.50 ± 5.07	63.25 ± 4.40	63.25 ± 4.62

Rezultati prikazani u Tabeli 1: pokazuju da modeli zasnovani na *LLM*, u zavisnosti od broja uključenih kriterijuma, postižu performanse koje su uporedive sa *DIDEX* modelom i standardnim modelima mašinskog učenja, uključujući stablo odlučivanja, nasumične šume i logističku regresiju. Najbolji *LLM* model (sa sedam kriterijuma) postiže tačnost od 65.62% ± 4.43%, što je veoma blizu rezultatu *DIDEX* modela (65.73% ± 12.16%) i nasumične šume (66.27% ± 5.39%).

Dodatno, primećujemo da se sa povećanjem broja kriterijuma u *LLM* modelima smanjuje standardna devijacija svih evaluacionih metrika, što ukazuje na stabilnije ponašanje modela pri predikciji. Ovo sugeriše da veći broj pažljivo odabranih atributa omogućava *LLM* modelu da bolje generalizuje i pravi konzistentnije predikcije, čak i kada performanse po proseku ostanu slične.

5. ZAKLJUČAK

U ovom radu predstavljen je metodološki okvir za integraciju velikih jezičkih modela (*LLM*) u proces konstrukcije *DEX* modela za podršku odlučivanju. Korišćenjem *LLM*-a za generisanje hijerarhije kriterijuma i objašnjenja, a zatim podataka za izgradnju tabela pravila odlučivanja, dobijen je interpretabilan, monotono-konzistentan model sa rezultatima koji su uporedivi sa savremenim pristupima mašinskog učenja. Osim što omogućava znatno bržu izradu modela u odnosu na prethodno razvijeni *DIDEX* pristup, predloženi pristup obezbeđuje veću transparentnost i lakšu komunikaciju između modela i krajnjih korisnika kroz objašnjenja koja se

dobijaju iz LLM-a, čime zadovoljava osnovne zahteve kvaliteta odlučivanja (5C princip).

Kao pravac za budući rad, predlaže se razvoj sistema zasnovanog na više LLM agenata koji bi predstavljali različite aktere u procesu donošenja odluka (eng. *multi-actor decision making*). Svaki agent bi bio osposobljen da sagleda problem iz sopstvenog ugla, uz sopstvene preferencije, ograničenja i domenske pretpostavke. Kroz dijalog, razmenu argumenata i upotrebu metodologija društvenog izbora (eng. *social choice theory*), kao što su pregovaračke igre (Kramár et al., 2022), deliberacija ili (samo-)refleksivno usklađivanje mišljenja, agenti bi mogli kolektivno da dođu do kompromisnog, reprezentativnog modela odlučivanja. Ovaj pristup otvara mogućnosti za automatizaciju kompleksnih procesa odlučivanja u kontekstima sa više aktera, kao što su javne politike, zdravlje ili strateško planiranje.

LITERATURA

- Akerkar, R., & Sajja, P. (2009). *Knowledge-based systems*. Jones & Bartlett Publishers.
- Barros, S. (2025). I Think, Therefore I Hallucinate: Minds, Machines, and the Art of Being Wrong. *arXiv preprint arXiv:2503.05806*.
- Bazerman, M. H., & Moore, D. A. (2012). *Judgment in managerial decision making*. John Wiley & Sons.
- Bohanec, M. (2021). From data and models to decision support systems: Lessons and advice for the future. In *EURO Working Group on DSS: A tour of the DSS developments over the last 30 years* (pp. 191–211). Springer.
- Bohanec, M. (2022). DEX (Decision EXpert): A qualitative hierarchical multi-criteria method. In *Multiple Criteria Decision Making: Techniques, Analysis and Applications* (pp. 39–78). Springer.
- Burstein, F., W Holsapple, C., & Power, D. J. (2008). Decision support systems: a historical overview. *Handbook on decision support systems 1: Basic themes*, 121–140.
- Carus, L., & Castillo, I. (2021). Managing risk in ski resorts: Environmental factors affecting actual and estimated speed on signposted groomed slopes in a cohort of adult recreational alpine skiers. *Plos one*, 16(8), e0256349.
- Delibašić, B., Radovanović, S., Jovanović, M. Z., Obradović, Z., Suknović, M., & Lojić, R. (2024). A study on ski groups size and their relationship to the risk of injury. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*, 238(4), 370–381.
- Delibašić, B., Radovanović, S., & Vukanović, S. (2023). A Decision Support System for Internal Migration Policy-Making.
- He, Q., Wang, Y., & Wang, W. (2024). Can Language Models Act as Knowledge Bases at Scale? *arXiv preprint arXiv:2402.14273*.
- Jones, C. R., & Bergen, B. K. (2025). Large language models pass the turing test. *arXiv preprint arXiv:2503.23674*.
- Kramár, J., Eccles, T., Gemp, I., Tacchetti, A., McKee, K. R., Malinowski, M., Graepel, T., & Bachrach, Y. (2022). Negotiation and honesty in artificial intelligence methods for the board game of Diplomacy. *Nature Communications*, 13(1), 7214.
- Meng, K., Bau, D., Andonian, A., & Belinkov, Y. (2022). Locating and editing factual associations in gpt. *Advances in neural information processing systems*, 35, 17359–17372.
- Moscatti, I. (2021). History of utility theory. *The Routledge Handbook of the Philosophy of Economics*, 23–36.
- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2023). A comprehensive overview of large language models. *arXiv preprint arXiv:2307.06435*.
- Radovanović, S., Bohanec, M., & Delibašić, B. (2023). Extracting decision models for ski injury prediction from data. *International Transactions in Operational Research*, 30(6), 3429–3454.
- Radovanović, S., Delibašić, B., & Vukanović, S. (2024). Combining LLM and DIDEX method to predict Internal Migrations in Serbia. *Human-Centric Decision and Negotiation Support for Societal Transitions*.
- Rink, P., & Brannath, W. (2025). Post-selection confidence bounds for prediction performance. *Machine Learning*, 114(3), 82.
- Turban, E., Sharda, R., Delen, D., Aronson, J., Lian, T.-P., & King, D. (2015). *Business intelligence and analytics: systems for decision support*. Pearson.