

Obnašanje velikih jezikovnih modelov na napovednih trgih: primerjava z odločitvenimi vzorci ljudi

Marjan Meglen

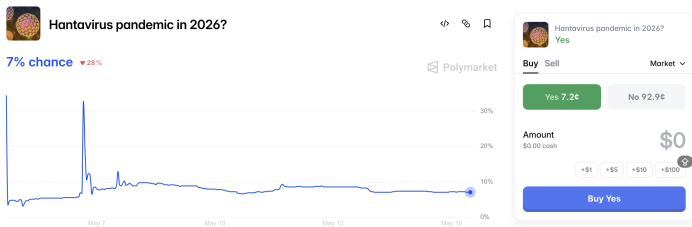
Mentor: dr. Aleksandar Tošić

Somentor: dr. Niki Hrovatin

2026

Napovedni trgi

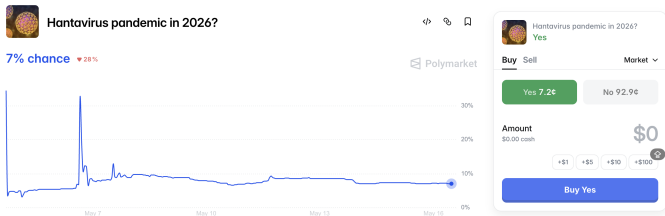
- ▶ Napovedni trgi so trgi, na katerih udeleženci trgujejo z izidi prihodnjih dogodkov.
- ▶ Pri binarnem trgu lahko udeleženec kupi pogodbo za izid YES ali NO.
- ▶ Cena YES delnice in cena NO delnice seštejeta v 1 dolar.



Slika: Primer binarnega trga na platformi Polymarket.

Interpretacija cen kot verjetnosti

- ▶ Cene teh delnic lahko interpretiramo kot kolektivno oceno udeležencev o tem, kako verjeten je določen razplet dogodka.
- ▶ Na dan 16. maj 2026 so udeleženci menili, da je verjetnost Hantavirus pandemije 7%.



Slika: Gibanje tržne verjetnosti na primeru Polymarket trga.

LLM in cene na trgih

- ▶ Pojavi vprašanje, kako takšen tržni signal kot je cena na trgu interpretira veliki jezikovni model.
- ▶ Zanj lahko predstavlja zgolj številčni podatek, lahko pa jo interpretira tudi kot oceno, ki so jo podali drugi udeleženci trga.
- ▶ Kako napoveduje verjetnost dogodka LLM, če mu podamo informacijo o preostalih udeležencih trga.

Motivacija raziskave

- ▶ Izhodišče predstavlja članek Kim: *LLMs Position Themselves as More Rational Than Humans*.
- ▶ Avtor preučuje vedenje LLM-jev v igrah iz teorije iger.
- ▶ Rezultati kažejo, da modeli pogosto razlikujejo med človeškimi in umetnimi akterji.
- ▶ V moji raziskavi je podobno vprašanje preneseno v kontekst napovednih trgov.

Osrednja ideja

Ali LLM drugače napoveduje izid trga, če so udeleženci ki so oblikovali verjetnost dogodka, predstavljeni kot ljudje, LLM agenti ali sploh niso omenjeni?

Hipoteze

- ▶ LLM bo v promptu dobival različne informacije.
- ▶ H_01 : Prikaz informacije o drugih udeležencih trga nima statistično značilnega vpliva na napovedi LLM modela.
- ▶ H_02 : Prikaz trenutnih tržnih verjetnosti nima statistično značilnega vpliva na napovedi LLM modela.
- ▶ H_03 : Prikaz novic nima statistično značilnega vpliva na napovedi LLM modela.

Pregled metodologije

- ▶ Metodologija je sestavljena iz treh glavnih delov:
 1. pridobivanje podatkov,
 2. izvedba eksperimenta z LLM modelom,
 3. statistična analiza rezultatov.
- ▶ Cilj metodologije je primerjati napovedi modela pri različnih informacijah v pozivu.

Pridobivanje podatkov: izbor trgov

- ▶ Na platformi Polymarket sem izbral 10 trgov iz leta 2025 z največjim obsegom vloženega denarja.
- ▶ Nato sem izločil športne dogodke, ker zaradi narave izbranih novic niso bili najbolj primerni za analizo.
- ▶ Po tem koraku je ostalo 6 trgov, večinoma politične narave.
- ▶ Primera trgov:
 - ▶ ali bo Donald Trump inavguriran kot predsednik ZDA,
 - ▶ odločitev ameriške centralne banke o obrestnih merah v oktobru 2025.

Pridobivanje podatkov: dnevne verjetnosti

- ▶ Za vsak izbrani trg sem pridobil dnevne verjetnosti za obdobje 20 dni pred razrešitvijo trga.
- ▶ Za vsak dan sta bili shranjeni tržni verjetnosti za izida YES in NO.

Datum	Verjetnost YES	Verjetnost NO
2025-01-01	p_{yes}	p_{no}
2025-01-02	p_{yes}	p_{no}
...
2025-01-20	p_{yes}	p_{no}

Tabela: Struktura dnevnih tržnih verjetnosti za posamezen trg.

Pridobivanje podatkov: novice

- ▶ Za isto 20-dnevno obdobje sem zbral novice iz štirih virov:
 - ▶ CNN,
 - ▶ Fox News,
 - ▶ MS Now,
 - ▶ NY Post.
- ▶ Ker je bilo člankov veliko, sem za vsak trg in vsak dan obdržal največ 5 najbolj relevantnih člankov iz vsakega vira.
- ▶ To pomeni največ 20 člankov (5x CNN, 5x Fox News, 5x MS Now, 5x NYPost) na dan oziroma največ 400 člankov za posamezen trg.

Izvedba eksperimenta

- ▶ Za vsak trg sem pripravil več konfiguracij poziva.
- ▶ Konfiguracije določajo:
 - ▶ ali model vidi trenutne tržne verjetnosti,
 - ▶ ali model vidi novice,
 - ▶ kako so predstavljeni drugi udeleženci trga.
- ▶ Za vsak dan v opazovanem obdobju se sestavi nov poziv za LLM model.
- ▶ Poziv vsebuje tržno vprašanje, trenutni datum, datum razrešitve in informacije, ki jih dovoljuje izbrana konfiguracija.

Konfiguracije poziva

Konf.	Tržne verjetnosti	Novice	Drugi udeleženci
1	da	da	LLM agenti
2	da	ne	LLM agenti
3	da	da	ljudje
4	da	ne	ljudje
5	da	da	brez navedbe
6	da	ne	brez navedbe
7	ne	da	LLM agenti
8	ne	ne	LLM agenti
9	ne	da	ljudje
10	ne	ne	ljudje
11	ne	da	brez navedbe
12	ne	ne	brez navedbe

Tabela: Vseh 12 konfiguracij poziva.

Primer poziva

Market question:

Will the Federal Reserve cut interest rates?

Current date: ...

Resolution date: ...

Other market participants are human traders.

Current market probabilities:

YES: 0.72

NO: 0.28

Relevant news:

...

Return a valid JSON object with:

reasoning and probability_yes.

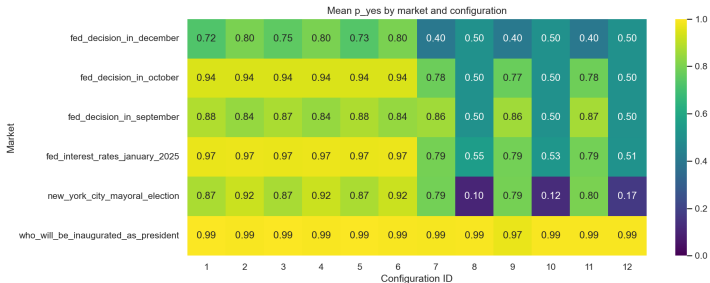
Analiza rezultatov

- ▶ Model za vsak poziv vrne kratko utemeljitev in oceno verjetnosti, da se trg razreši z izidom YES.
- ▶ Za analizo rezultatov sem izračunal opisne statistike po trgih in konfiguracijah.
- ▶ Pripravljeni so bili grafični prikazi gibanja napovedi LLM modela skozi čas.
- ▶ Za preverjanje hipotez sem uporabil parne t-teste.

Primerljivost pogojev

Primerjave se izvajajo znotraj istega trga in istega datuma, da so pogoji med konfiguracijami primerljivi.

Pregled po konfiguracijah



Slika: Povprečni p_{yes} po trgih in konfiguracijah.

- ▶ Napovedi se razlikujejo med trgi in med konfiguracijami.
- ▶ Najbolj izstopa razlika med konfiguracijami, kjer so tržne verjetnosti prikazane, in tistimi, kjer niso.

Vpliv vrste drugih udeležencev

- ▶ To je bil osrednji del raziskave.
- ▶ Primerjave so izolirale samo razliko med opisom drugih udeležencev.
- ▶ Rezultati kažejo šibek in nekonsistenten učinek.

Mera	Rezultat
Število parnih primerjav	72
Statistično značilne primerjave	3
Mediana absolutne razlike	0,001
Povprečna absolutna razlika	0,006

Tabela: Povzetek učinka vrste drugih udeležencev.

Značilne primerjave za druge udeležence

Trg	Primerjava	Δp_{yes}
December Fed	LLM – ljudje	-0,026
September Fed	LLM – ljudje	0,007
January rates	LLM – brez navedbe	0,047

Tabela: Statistično značilne primerjave za vrsto drugih udeležencev.

Interpretacija

Smer učinka ni stabilna: v enem primeru so napovedi pri LLM udeležencih nižje, v drugem višje.

Vpliv tržnih verjetnosti

- ▶ Najmočnejši in najbolj konsistenten učinek v eksperimentu.
- ▶ Ko model vidi tržne verjetnosti, se njegove napovedi pogosto približajo tržnemu signalu.
- ▶ Ta signal lahko zasenči šibkejše informacije, kot je opis drugih udeležencev.

Mera	Rezultat
Število parnih primerjav	36
Statistično značilne primerjave	29
Mediana absolutne razlike	0,241
Povprečna absolutna razlika	0,259

Tabela: Povzetek učinka prikaza tržnih verjetnosti.

Vpliv novic

- ▶ Novice imajo izrazit, vendar manj enoten učinek kot tržne verjetnosti.
- ▶ Pri nekaterih trgih novice zvišajo p_{yes} , pri drugih ga znižajo.
- ▶ Učinek je posebej pomemben, ko model ne vidi tržnih verjetnosti.

Mera	Rezultat
Število parnih primerjav	36
Statistično značilne primerjave	30
Mediana absolutne razlike	0,051
Povprečna absolutna razlika	0,151

Tabela: Povzetek učinka prikaza novic.

Kaj pomenijo rezultati?

- ▶ Model se očitno odziva na informacije v pozivu, vendar ne enako močno na vse informacije.
- ▶ Tržne verjetnosti delujejo kot močan sidrni signal.
- ▶ Novice delujejo kot vsebinski signal, zato je smer učinka odvisna od konkretnega trga.
- ▶ Vrsta drugih udeležencev ima šibkejši vpliv, kot bi pričakovali iz motivacije po Kimovem članku.

Povezava s Kimovim člankom

- ▶ Kim pokaže, da LLM-ji v strateških igrah razlikujejo med tipi akterjev.
- ▶ V tej raziskavi se ta motivacija prenese v napovedne trge.
- ▶ Napovedni trg ni neposredna strateška igra.
- ▶ Model tu ocenjuje prihodnji dogodek, ne izbira poteze proti nasprotniku.

Glavna razlaga

Informacija o drugih udeležencih je lahko relevantna, vendar jo v tej zasnovi pogosto preglasijo tržne verjetnosti in novice.

Omejitve

- ▶ Uporabljen je bil omejen nabor šestih trgov.
- ▶ Analiziran je bil en velik jezikovni model.
- ▶ Uporabljena je bila ena osnovna oblika poziva.
- ▶ Testi so bili izvedeni ločeno po trgih in konfiguracijah.
- ▶ Rezultatov zato ni mogoče neposredno posplošiti na vse LLM modele ali vse napovedne trge.

Možnosti za prihodnje delo

- ▶ Uporaba več različnih LLM modelov.
- ▶ Več trgov in več ponovitev za vsako konfiguracijo.
- ▶ Dodatne formulacije stavkov o drugih udeležencih.
- ▶ Analiza interakcij med tržnimi verjetnostmi, novicami in tipom udeležencev.
- ▶ Primerjava s človeškimi napovedovalci ali dejanskimi tržnimi cenami.





Zaključek

- ▶ Prikaz tržnih verjetnosti ima najmočnejši in najbolj konsistenten vpliv na napovedi modela.
- ▶ Novice imajo pomemben, vendar tržno specifičen vpliv.
- ▶ Vrsta drugih udeležencev ima šibek in nekonsistenten učinek.
- ▶ Motivacija iz Kimovega članka se v napovednih trgih pokaže le delno.

Glavna ugotovitev

LLM napovedi so najbolj občutljive na neposredne informacije o trgu in dogodku; informacija o tem, kdo so drugi udeleženci, v tej zasnovi ne prevlada nad številčnimi in vsebinskimi signali.

Literatura I

-  K.-H. Kim, “LLMs position themselves as more rational than humans: Emergence of ai self-awareness measured through game theory,” *arXiv preprint arXiv:2511.00926*, 2025.
-  J. Wolfers and E. Zitzewitz, “Prediction markets,” *Journal of Economic Perspectives*, vol. 18, no. 2, pp. 107–126, 2004.
-  J. Wolfers and E. Zitzewitz, “Interpreting prediction market prices as probabilities,” Tech. Rep. 12200, National Bureau of Economic Research, 2006.
-  N. Lorè and B. Heydari, “Strategic behavior of large language models and the role of game structure versus contextual framing,” *Scientific Reports*, vol. 14, p. 18490, 2024.

Literatura II



Polymarket, “Historical timeseries data.” <https://docs.polymarket.com/developers/CLOB/timeseries,2026>.