

Kako nas oblikujejo algoritmi priporočil na družbenih omrežjih?



■ **Octavian M. Machidon**, dr. znanosti s področja elektrotehnike in telekomunikacij ter diplomirani teolog, je docent na Fakulteti za računalništvo in elektrotehniko Univerze v Ljubljani.

Kdo odloča, katere vsebine si ogledujete na TikToku, YouTubu ali Instagramu? Ste prepričani, da ste to vi? Zmogljivi algoritmi priporočil v ozadju vplivajo na vsak vaš premik, vsehček ali ogled videa. Ti sistemi so zasnovani tako, da prilagodijo vašo izkušnjo, vendar pri tem oblikujejo veliko več kot le vaš tok objav na družbenih omrežjih – vplivajo na vaše misli, vedenje in celo na vaše odločitve.

Algoritmi

Algoritmi priporočil so ena izmed vrst umetne inteligence, s katerimi se ljudje najpogosteje srečujejo. Postali so nepogrešljiv del našega digitalnega vsakdana, saj nam pomagajo krmariti po poplavi vsebin, ki so na voljo na spletu. Razvili so se kot rešitev za preobremenjenost z informacijami in so nadomestili starejše pristope, kot sta kronološki prikaz objav ali prikaz objav glede na omrežje, ki ne zmoreta več slediti hitri rasti digitalnih vsebin (Narayanan, 2023). Danes ti algoritmi poganjajo vse od TikToka strani »For You«, predlogov videov na YouTubu, priporočil filmov na Netflixu do prilagojenih novičarskih tokov na Googlu. Njihov cilj je preprost, a zelo vpliven: pritegniti vašo pozornost, vas obdržati na platformi in pogosto ustvariti prihodek z oglasi.

Ti sistemi delujejo tako, da analizirajo ogromne količine podatkov – vaše vedenje, preference in celo to, kaj so všečkali ali gledali uporabniki, ki so vam podobni. Na primer: TikTakov algoritem ne priporoča naključnih videov; prilagodi vaš tok vsebin glede na to, kaj je všeč ljudem, ki so v podobnem kontekstu kot vi. Tako ustvarja izjemno privlačno – nekateri bi rekli tudi zasvojljivo – izkušnjo.

Razumevanje, kako ti algoritmi delujejo, je ključno za učitelje, starše in skrbnike. Algoritmi priporočil ne oblikujejo le tega,

kar otroci in najstniki vidijo na spletu – oblikujejo tudi njihov pogled na svet. V nadaljevanju bomo raziskali njihov vpliv na mlade, etične pomisleke, ki jih povzročajo, in kaj lahko storimo, da zmanjšamo njihova tveganja.

Algoritmi priporočil in njihovo delovanje

Sistemi priporočil ta izziv rešujejo z osebnimi predlogi, katerih cilj je pritegniti zanimanje uporabnikov, ohraniti njihovo pozornost in graditi zvestobo platformi. Čeprav želijo izboljšati 'zadovoljstvo' uporabnikov, imajo pomembno vlogo tudi pri ustvarjanju prihodkov, pogosto s pomočjo ciljnega oglaševanja.

V svojem jedru skušajo ti algoritmi odgovoriti na preprosto vprašanje: kako so se uporabniki s podobnimi interesi odzvali na podobne vsebine? To dosežejo z analizo različnih vrst podatkov: vzorcev vedenja uporabnikov, povezav z drugimi uporabniki na družbenih omrežjih in demografskih značilnosti. Na podlagi prepoznavanja trendov in podobnosti oblikujejo osebne izkušnje, ki se tesno ujemajo z interesi posameznikov.

Algoritmi priporočil večinoma delujejo po dveh glavnih pristopih: sodelovalno filtriranje (angl. *collaborative filtering*) in filtriranje na podlagi vsebine (angl.

content-based filtering). Sodelovalno filtriranje analizira vedenje uporabnikov s podobnimi profili. Na primer, če Nina, najstnica iz Maribora, uporablja TikTok, algoritem platforme preučuje, katere vsebine so privlačne drugim najstnicam v mestnih območjih Slovenije, kaj jih najbolj pritegne in ohrani na aplikaciji. Na podlagi teh podatkov Ninina stran prejme podobne predloge, saj algoritem predvideva, da bodo tudi njej všeč. Po drugi strani pa filtriranje na podlagi vsebine ocenjuje značilnosti videov, ki jih je Nina že gledala. Če pogosto gleda videe z izzivi plesa, ji bo algoritem predlagal več takšnih vsebin, da bi ohranil njeno zanimanje.

Sodobni algoritmi pogosto združujejo oba pristopa v hibridne sisteme, ki jih poganja strojno učenje, da bi izboljšali natančnost in ustreznost predlogov. Njihov glavni cilj je preoblikovati vedenje uporabnikov v 'berljive' podatke, s čimer ustvarijo algoritemski profil za vsakega uporabnika. Ti profili – sestavljeni iz merljivih podatkovnih točk, kot so zgodovina ogledov, preference in interakcije – so osnova za prilagajanje vsebin, ki se ujemajo z individualnimi zanimanji. Tako algoritmi priporočil močno vplivajo na to, kaj uporabniki vidijo, s čim se ukvarjajo in kako na koncu dojemajo ter uporabljajo digitalni svet.

Pomembne skrbi glede algoritmov priporočil

Čeprav algoritmi priporočil optimizirajo uporabniško izkušnjo in angažiranost, povzročajo pomembne etične skrbi, saj temeljijo na algoritmičnih interpretacijah vedenja uporabnikov. Pogosto izkoriščajo individualne ranljivosti, kot so zasvojenost, motnje hranjenja, težave s samopodobo, anksioznost in depresija, z namenom maksimiranja angažiranosti. Če na

primer Nina, naša najstnica iz Maribora, občuti negotovost glede svojega videza in na TikToku začne iskati vsebine, povezane s telesno podobo ali dietami, bo algoritem zaznal to vedenje. Da bi jo obdržal na platformi, ji bo začel prikazovati še več takšnih vsebin, s čimer bo njeno negotovost dodatno okrepil in ji vsilil škodljive vzorce. Takšen pristop lahko uporabnike, kot je Nina, ujame v začaran krog, kjer so nenehno izpostavljeni vse ožjim in pogosto nezdravim vsebinam. To pogosto vodi v tako imenovano »doomscrolling« – pretirano porabo vsebin, povezanih s samopoškodbami, dietami ali idealiziranimi telesnimi podobami, kar še pogloblja njihove ranljivosti (Panoptikon Foundation, 2023; Henry idr., 2023).

Ta vzorec povečevanja ranljivosti ima resnične posledice, še posebej za mlade. Tragičen primer je Molly Russell, 14-letnica, ki si je leta 2017 vzela življenje, potem ko je bila na platformah, kot je Instagram, izpostavljena vsebinam, povezanim s samopoškodbami in samomorom. Algoritmi, zasnovani za maksimiranje angažiranosti, so Molly ujeli v škodljiv krog vse bolj vznemirjajočih vsebin. Njena zgodba je postala globalni poziv k ukrepanju, ki je spodbudil

učitelje, starše in odločevalce, da ponovno preučijo etiko teh sistemov in zahtevajo boljše zaščitne ukrepe (Djeffal idr., 2021). Na žalost Mollyjin primer ni osamljen. Mnogi najstniki so doživeli podobne usode zaradi škodljivih vsebin, ki so jih okrepili algoritmi družbenih omrežij. Nevladne organizacije, kot je ParentsSOS: Parents for Safe Online Spaces, so dokumentirale številne primere, v katerih so mladi izgubili življenje – bodisi zaradi samomora bodisi nesreč – potem ko so bili izpostavljeni nevarnim spletnim vsebinam. Te organizacije si prizadevajo ozaveščati javnost z deljenjem zgodb otrok in najstnikov, katerih ranljivosti so te platforme izkoriščale, ter zagovarjati močnejše zaščitne ukrepe za preprečevanje nadaljnjih tragedij (ParentsSOS, 2024).

Odsev nujnosti tega problema je tudi tožba mesta New York, vložena v začetku leta 2024, proti družbenim omrežjem, kot so TikTok, Instagram in YouTube. Tožba obtožuje platforme, da izkoriščajo duševno zdravje mladih uporabnikov preko škodljivih vsebin, ki jih algoritmi priporočil dodatno spodbujajo. Povod za tožbo je bila zaskrbljujoča situacija, ko so bile ustanove za duševno zdravje v državi preplavljene

ne z otroki in mladimi, ki so trpeli zaradi težav, povezanih z uporabo teh platform (Murphy Kelly, 2024).

Vse družbene platforme se soočajo z enakimi težavami

Študija *Amnesty International*, izvedena leta 2023 v sodelovanju z Algorithmic Transparency Institute in AI Forensics, je razkrila zaskrbljujoče učinke TikTovih algoritmov priporočil na otroke in mlade (Amnesty International, 2023). Uporabniki, ki so kazali zanimanje za teme o duševnem zdravju, so po nekaj urah uporabe prejeli skoraj polovico videov, ki so bili ocenjeni kot škodljivi za duševno počutje – desetkrat več kot uporabniki, ki teh interesov niso izkazovali. Veliko teh videov je normaliziralo ali celo povečevalo škodljiva vedenja, kot so samopoškodbe in samomor. Študija je prav tako kritizirala TikTokovo zasnovano, usmerjeno v čim večjo angažiranost, saj ta nesorazmerno spodbuja tovrstne vsebine in uporabnike zadržuje na platformi dlje časa, kar spodbuja zasvojenost in čustvene stiske. Toda ta problem ni omejen le na TikTok. Tudi YouTubeov sistem priporočil je bil



Foto: Nataša Pezdir



Foto: Benedikt Lavrih

deležen kritik zaradi usmerjanja uporabnikov k škodljivim vsebinam. Leta 2019 je organizacija *Mozilla* začela globalno kampanjo, v kateri je skovala izraz »YouTube obžalovanja« (angl. YouTube regrets), s katerim je opisala, kako algoritmi uporabnike vodijo k zavajajočim ali škodljivim vsebinam, vključno s teorijami zarote in ekstremističnimi materiali. Med letoma 2020 in 2021 je Mozillin pripomoček RegretsReporter razkril, da je 71 % videov, ki so jih uporabniki označili kot obžalovanja vredne, izviralo iz YouTubeovih priporočil, pri čemer so priporočila imela 40 % večjo verjetnost, da bodo povzročila negativne izkušnje v primerjavi z iskalnimi rezultati. Poleg tega je bilo veliko teh videov nepovezanih s prejšnjo aktivnostjo uporabnikov, znaten delež pa je celo kršil YouTubeove smernice skupnosti (McCrosky in Geurkink, 2021).

Ti primeri kažejo zaskrbljujoč vzorec na vseh platformah: algoritmi, zasnovani za optimizacijo angažiranosti, pogosto izkoriščajo vedenje uporabnikov na način, ki spodbuja zasvojenost in krepi škodljive vsebine.

Ustvarjanje varnejšega digitalnega okolja

Za učitelje in starše so te teme izjemno pomembne, saj številni mladi vsak dan preživijo ure na platformah, kot so YouTube, TikTok in Instagram. Raziskovalno novinarstvo je dodatno osvetlilo ta problem. V letih 2022 in 2023 je Bloomberg poročal o škodljivih učinkih TikTokovih algoritmov priporočil, ki so bili povezani s primeri sa-

momorov med mladostniki in nenamernimi smrtnimi primeri. Te preiskave so pokazale, kako algoritmi ranljive uporabnike potiskajo proti nevarnim vsebinam, kar poudarja nujno potrebo po spremembah (Carville, 2022 in 2023).

Še posebej zaskrbljujoče je pomanjkanje mehanizmov za preprečevanje škode. Trenutni sistemi priporočil nimajo 'izklopne tipke' in ne sprožijo opozoril, da bi zaščitili uporabnike, kot je Nina – naša najstnica iz primera – ali druge ranljive mlade pred začaranimi krogi škodljive uporabe vsebin. Učitelji, starši in skrbniki morajo prepoznati to pomanjkljivost in sodelovati pri ustvarjanju varnejšega digitalnega okolja. Šole in družine imajo ključno vlogo pri poučevanju mladih, da prepoznajo škodljive vzorce, kritično razmišljajo o tem, kar vidijo na spletu, in odgovorno uporabljajo digitalne vsebine.

Ta skrb je zelo aktualna tudi v Sloveniji, kjer so družbena omrežja izjemno priljubljena. TikTok na primer uporablja več kot 600.000 Slovencev, pri čemer prevladujejo mlajše generacije. Devet od desetih slovenskih uporabnikov TikToka dostopa do platforme večkrat na dan, generacija Z pa dnevno porabi več kot dve uri za ogled videov (Mediana, 2024). Visoka angažiranost na tej platformi in njen vpliv na oblikovanje trendov poudarjata potrebo po večji ozaveščenosti o tveganjih, ki jih algoritmi priporočil predstavljajo za mlado, dojemljivo občinstvo.

Za spopadanje s temi izzivi bi morale šole vključiti digitalno pismenost in izobraževanje o varnosti na spletu v svoje učne načrte, da bi učence opremile z orodji

za odgovorno navigacijo po družbenih omrežjih. To mora dopolnjevati ozaveščanje staršev, da razumejo tveganja, povezana z uporabo družbenih omrežij, in da lahko vodijo svoje otroke k bolj zdravim digitalnim navadam. Skupaj moramo stremeti k ustvarjanju varnejše in premišljene digitalne prihodnosti za naslednjo generacijo.

Seveda je ključen vidik te težave dostop otrok in mladostnikov do pametnih telefonov in družbenih omrežij. Upravljanje tega dostopa je bistveni korak pri reševanju širših vprašanj, ki jih te tehnologije povzročajo. Ta tema pa si zasluži poglobljeno obravnavo, ki bo osrednja tema našega naslednjega članka. ◀

Viri

- Amnesty International (2023): Driven into darkness: How TikTok's 'For You' feed encourages self-harm and suicidal ideation. *Technical report*, 7. 11. 2023. Pridobljeno 20. 11. 2024 s spletne strani: <https://www.amnesty.org/en/documents/pol40/7350/2023/en/>.
- Carville, Olivia (2022): TikTok's viral challenges keep luring young kids to their deaths. *Bloomberg Businessweek*, 30. 11. 2022. Pridobljeno 20. 11. 2024 s spletne strani: <https://www.bloomberg.com/news/features/2022-11-30/is-tiktok-responsible-if-kids-die-doing-dangerous-viral-challenges>.
- Carville, Olivia (2023): TikTok's algorithm keeps pushing suicide to vulnerable kids. *Bloomberg Businessweek*, 20. 4. 2023. Pridobljeno 20. 11. 2024 s spletne strani: <https://www.bloomberg.com/news/features/2023-04-20/tiktok-effects-on-mental-health-in-focus-after-teen-suicide>.
- Djeflal, Christian; Hitrova, Christina; Magrani, Eduardo (2021): Recommender systems and autonomy: A role for regulation of design, rights, and transparency. *Indian Journal of Law & Technology*, 17(1).
- Henry, Nathan; Pedersen, Mangor; Martin, Jamin idr. (2023): Reducing echo chamber effects: An allostatic regulator for recommendation algorithms. *ResearchSquare*. Pridobljeno 13. 11. 2024 s spletne strani: <https://www.researchsquare.com/article/rs-3708350/v1>.
- McCrosky, J.; Geurkink, B. (2021): *YouTube Regrets: A Crowdsourced Investigation into YouTube's Recommendation Algorithm*. Mozilla Foundation. Pridobljeno 13. 11. 2024 s spletne strani: https://assets.mofoprod.net/network/documents/Mozilla_YouTube_Regrets_Report.pdf.
- Mediana (2024): *TikTok v Sloveniji najbolj priljubljen med generacijo Z, z njim pa se spogledujejo tudi druge generacije*. Pridobljeno 20. 11. 2024 s spletne strani: <https://mediana.si/tiktok-v-sloveniji-najbolj-priljubljen-med-generacijo-z-z-njim-pa-se-spogledujejo-tudi-druge-generacije/>.
- Murphy Kelly, Samantha (2024): New York City sues social media platforms over youth mental health crisis. *CNN Business*, 16. 2. 2024. Pridobljeno 20. 11. 2024 s spletne strani: <https://edition.cnn.com/2024/02/14/tech/new-york-city-sues-social-media-platforms-youth-mental-health/index.html>.
- Narayanan, Arvind (2023): Understanding social media recommendation algorithms. *Knight First Amendment Institute at Columbia University*, 9. 3. 2023. Pridobljeno 13. 11. 2024 s spletne strani: <https://knightcolumbia.org/content/understanding-social-media-recommendation-algorithms>.
- Panoptikon Foundation (2023): *Fixing recommender systems – briefing note*. Technical report. Irish Council for Civil Liberties and the Panoptikon Foundation. Pridobljeno 13. 11. 2024 s spletne strani: <https://en.panoptikon.org/fixing-recommender-systems-identification-risk-factors-meaningful-transparency-and-mitigation>.
- ParentsSOS (2024): *Parents for Safe Online Spaces – Advocating for Online Harms Prevention*. Pridobljeno 20. 11. 2024 s spletne strani: <https://www.parentssos.org>.