
Veštačka inteligencija (AI) u funkciji prepoznavanja emocija i nasilničkog mentaliteta

Momčilo B. Bajac¹ i Željko Đ. Bjelajac²

¹Univerzitet UNION Nikola Tesla u Beogradu
Fakultet za menadžment, Sremski Karlovci

²Univerzitet Privredna akademija u Novom Sadu
Pravni fakultet za privredu i pravosuđe u Novom Sadu

Informacije o članku*

Originalni naučni rad • UDK: 004.8:343.22

Volume: 19 Broj: 2, stranice: 277–297.

Primljeno: 28. maj 2022 • Prihvaćeno: 12. jun 2022.

<https://doi.org/10.51738/Kpolisa2022.19.2p.277bb>

Podaci o autorima

Momčilo B. Bajac  <https://orcid.org/0000-0003-2115-6373>

Željko Đ. Bjelajac  <https://orcid.org/0000-0003-4953-8779>

Ne postoji konflikt interesa u pogledu izrade ovog teksta.

Autor za korespondenciju: Momčilo Bajac, Fakultet za menadžment,

Njegoševa 1, 21205 Sremski Karlovci, Srbija.

Email: momcilo.bajac@famns.edu.rs

* Cite (APA):

Bajac, M. B., & Bjelajac, Ž. Đ. (2022). Veštačka inteligencija (AI) u funkciji prepoznavanja emocija i nasilničkog mentaliteta [Artificial Intelligence (AI) in the Function of Recognizing Emotions and Violent Mentality]. *Kultura polisa*, 19(2), 277–297.

<https://doi.org/10.51738/Kpolisa2022.19.2p.277bb>

Sažetak

U ovom radu diskutujemo uticaj najnovijih tehnologija, prvenstveno veštačke inteligencije (AI) na šire društveno i prirodno okruženje, kao i na individualne živote ljudi u sve većem rasponu konteksta. Govorimo o AI kao kompleksnim ekosistemima napravljenim od prirodnih resursa, energije, infrastrukture, ljudskog rada, logistike, istorije, podataka i mnoštva klasifikacija. Uprkos tome što za sada AI nije ni veštačka a ni toliko inteligentna kao što se to predstavlja široj javnosti, postoji jaka težnja za njenom univerzalizacijom i apsolutizacijom. U ovom radu posebno smo diskutovali mogućnost primene AI u svrhu prepoznavanja emocija i nasilničkog mentaliteta u cilju prevencije nasilničkog ponašanja. To je zahtevalo analizu prakse prikupljanja podataka i njihovog klasifikovanja u fiksne kategorije kako bi se postigli “pouzdani” izlazi pretrage za različite potrebe. Profil nasilničke ličnosti je određen kako socijalno-relacionim tako i biološko-psihološkim karakteristikama, što predstavlja veliki broj varijabli koje treba uzeti u obzir prilikom konstruisanja algoritma za prepoznavanje emocija i nasilničke ličnosti. Naš zaključak je da sa trenutnim stepenom razvoja AI i veoma svedenim i diskretnim algoritmima nije moguće uzeti u obzir toliko podataka da bi se dobio pouzdan izlaz za predviđanje ponašanja kao i preduzimanje eventualnih mera, a da nisu diskriminatorne i opasne po pojedince i društvo u celini. Ali s obzirom na neverovatno brz tehnološki razvoj, naročito u oblasti kvantnog računarstva, pred ovom oblasti stoji veoma dinamična budućnost.

Ključne reči: algoritam, skupovi podataka, klasifikacija, facijalne ekspresije, nasilnička ličnost, kvantno računarstvo

Veštačka inteligencija (AI) u funkciji prepoznavanja emocija i nasilničkog mentaliteta

Kada govorimo o veštačkoj inteligenciji, prva asocijacija nam je proizvod visoke tehnologije, nešto nematerijalno, kodirano, što egzistira u virtualnom prostoru svuda oko nas. Međutim, AI je vrlo kompleksan ekosistem koji uključuje mnoštvo materijalnih i nematerijalnih aspekata i ima vrlo snažan impakt na naše ekonomsko, socijalno pa i političko okruženje.

Sam pojam inteligencije je veoma kompleksan i još nedovoljno istražen. Postoji veoma mnogo predrasuda i mitova oko inteligencije koji su imali veoma loše konsekvence kroz istoriju čovečanstva. Prvi mit polazi od stava da se uz dovoljnu obuku ili dovoljno resursa, ljudska inteligencija može stvoriti od nule, bez obraćanja pažnje na fundamentalne načine na koje su ljudi povezani i smešteni u široj konstelaciji raznolike stvarnosti. Drugi mit da je inteligencija nešto što je prirodno i nezavisno od društvenih, kulturnih, istorijskih i političkih sila. U stvari, koncept inteligencije je vekovima naneo ogromnu štetu i korišćen je da opravda odnose dominacije, od ropstva do eugenike. O kompleksnosti inteligencije koja je i danas predmet interesovanja istraživača iz skoro svih oblasti nauke, govori i činjenica o već razvijenim konceptima emocionalne, socijalne, moralne i duhovne inteligencije. Inteligencija predstavlja sposobnost snalaženja u novim, neočekivanim situacijama. Iz navedenog proizilazi da je suštinska karakteristika inteligencije uviđanje relevantnih odnosa u konkretnoj situaciji. Emocionalna inteligencija je sposobnost prepoznavanja osećanja, njihovog jasnog identifikovanja, razumevanja, sposobnosti kontrolisanja i korišćenja za izražavanje misli. To je sposobnost da se efikasno održava veza između emocija i mišljenja, da se upotrebe emocije kako bi se olakšalo rasuđivanje, da se inteligento rasuđuje o emocijama. Ljudi emocije retko izražavaju rečima. Emocije se mnogo češće izražavaju drugim znakovima. Sposobnost da se prepoznaju neizgovoreni znakovi je ključ za identifikaciju i razumevanje osećanja na osnovu izraza lica, pokreta ili tona glasa. Još je Aristotel rekao da se "svako može naljutiti – to je lako.

Ali naljutiti se na pravu osobu, do ispravnog stepena, u pravi trenutak, zbog ispravnog razloga, na ispravan način – to nije lako." Ako tome dodamo i duhovnu inteligenciju putem koje osoba promišlja o suštini sopstvenog postojanja i traga za odgovorima o pitanjima smisla, onda je veoma jasno koliko je kompleksan posao implementacije AI u sve sisteme koji se odnose na ljudsko ponašanje i emocionalne manifestacije koje ga prate.

Uprkos ovim saznanjima, postoje i značajno pojednostavljene ideje inteligencije, po kojima su ljudi puki sistem za obradu informacija, pa prema tome naučnici AI mogu da stvore mašinu koja uči kao što uči dete. Takve ideje i njihovi relikti su neretko ugrađeni u temelje današnjih koncepata veštačke inteligencije, koja sve više prožima svakodnevni život ljudi. Profesor filozofije Hubert Drajfus je u svom radu Šta kompjuteri ne mogu da urade, na takav pristup uzvratilo argumentom da mozak obrađuje informacije na potpuno drugačiji način od kompjutera. Naime, "ljudska inteligencija u velikoj meri uključuje mnoge nesvesne i podsvesne procese, dok računari zahtevaju da svi procesi i podaci budu eksplicitni i formalizovani" (Dreyfus, 1999). Kao rezultat toga, manje formalni aspekti inteligencije moraju biti apstrahovani, eliminisani ili aproksimirani za računarsku primenu, ostavljajući ih nesposobnim da obrađuju informacije o kompleksnim situacijama kao što to čine ljudi. Emocija je reakcija subjekta na stimulus koji je pojedinac ocenio kao važan, a koja priprema subjekta za adaptivnu aktivnost. Mi ono što saznajemo vrednujemo osećanjima, pri čemu emocionalne reakcije nisu odgovori na bilo kakve stimuluse iz spoljašnje sredine, već samo na one stimuluse koje ocenjujemo kao nama bitne.

Ponešeni naglim tehnološkim napretkom od kraja 60-tih godina, naučnici i vodeće IT korporacije polako "zaboravljaju" na rane diskusije oko toga šta AI može da uradi a šta ne, dok se skepticizam u odnosu na njene stvarne mogućnosti polako istopio. Od sredine 2000-ih, AI se brzo proširila kao polje u akademskim krugovima i industriji. Sada mali broj moćnih tehnoloških korporacija kao što su Alfabet, Google, Facebook, Amazon, Nvidia, razvija i primenjuje sisteme veštačke inteligencije na planetarnom

nivou, a njihovi sistemi se ponovo pozdravljaju kao uporedivi sa ljudskom inteligencijom (AGI – Artificial General Intelligence) ili čak superiorniji od ljudske inteligencije (ASI – Artificial Super Intelligence).

Kada govorimo o AI kao kompleksnim ekosistemima uzimamo u obzir činjenicu da je ona napravljena od “prirodnih resursa, goriva, ljudskog rada, infrastrukture, logistike, istorije i mnoštva klasifikacija” (Krawford, 2021). Današnji sistemi veštačke inteligencije, uprkos naglašavanju njenog značaja i mogućnosti, nisu autonomni, racionalni ili sposobni da razaznaju bilo šta bez opsežne, računarski intenzivne obuke sa velikim skupovima podataka sa unapred definisanim pravilima i mnoštvom drugih resursa neophodnih za njeno projektovanje. “U stvari, veštačka inteligencija kakvu poznajemo u potpunosti zavisi od mnogo šireg skupa političkih i društvenih struktura, a zbog kapitala potrebnog za izgradnju veštačke inteligencije u globalnim razmerama, AI sistemi su uglavnom dizajnirani da služe postojećim dominantnim interesima i manifestaciji njihove moći” (Krawford, 2021). Da bi u celini razumeli suštinu i kompleksnost AI, moramo uzeti u obzir države i korporacije koje njome upravljaju i dominiraju, ekstrakcijsko rudarenje koje ostavlja trag na planeti, masovno prikupljanje podataka radi nadzora (nadzorni kapitalizam) i duboko nejednake i sve eksploativne prakse rada koje ovaj sistem održavaju.

Uprkos tome što za sada AI nije ni veštačka a ni toliko inteligentna kao što se predstavlja široj javnosti, postoji jaka težnja kako iz naučnih tako i iz industrijskih krugova za njenom univerzalizacijom i apsolutizacijom. AI treba da stvori digitalni odraz stvarnog sveta koji će mapirati celokupni svet objekata, a na duge staze, veštačka inteligencija treba da postane jedina nauka sposobna da obavi bilo koji intelektualni zadatak. Sam pojam veštačke inteligencije je ambivalentan. Pojam veštačka inteligencija se većinom koristi kada se govori o masivnoj industrijskoj formaciji koja uključuje politiku, rad, kulturu i kapital, uslovno rečeno za “marketinške potrebe” prema široj javnost, dok se o mašinskom učenju govori u kontekstu tehničkih pristupa u njenom konstruisanju.

AI i resursi

Stvaranje savremenih sistema veštačke inteligencije zavisi od eksploatacije energije i mineralnih resursa sa planete, jeftine radne snage i podataka u velikim razmerama. Izraz „veštačka inteligencija“ može da se poziva na ideje algoritama, podataka i arhitekture klauda, ali ništa od toga ne može da funkcioniše bez resursa koji čine osnovne komponente računarstva. Kompletan lanac nabavke veštačke inteligencije zahteva ogromnu količinu kapitala, radne snage i resursa Zemlje. Masivni ekosistem veštačke inteligencije oslanja se na mnoge vrste ekstrakcije: od prikupljanja podataka dobijenih iz naših svakodnevnih aktivnosti i izražavanja, do iscrpljivanja prirodnih resursa i eksploatacije radne snage širom sveta kako bi se ova ogromna planetarna mreža mogla izgraditi i održavati.

AI automatizacija radnog mesta je već davno uspostavljena činjenica savremenog rada. Proizvodna montažna linija danas već ima svoje analoge i u uslužnim delatnostima, sektoru komunikacija, a takozvani radnici znanja za koje se pretpostavlja da su manje ugroženi automatizacijom, sve više su podvrgnuti nadzoru na radnom mestu, automatizaciji procesa i gubljenju razlike između radnog i slobodnog vremena. Uobičajeni argument za proširenje AI sistema i automatizacije procesa je da živimo u vremenu korisne saradnje između ljudi i AI. Ali ova saradnja nije rezultat “dogovora” već prisilnog angažovanja, gde se od radnika očekuje da se uvek iznova i iznova obučavaju, nastave i bespogovorno prihvate svaku novu tehničku inovaciju. I to je bolji ishod koji zaposleni mogu da očekuju. Jedan od najznačajnijih eksponenata Svetskog ekonomskog foruma (WEF), Juval Noa Harari, u svom obraćanju je konstatovao da u većini slučajeva, automatizacija poslova kroz implementaciju AI sistema znači gubljenje radnih mesta i stvaranja armije “nezapošljivih i bespotrebnih ljudi, jer kao što su se u prošlosti ljudi plašili eksploatacije, u 21. veku ljudi će se plašiti beskorisnosti. Mnogo je gore biti beskoristan nego eksploatisan”. Ovakva “ponuda” partnerstva čoveka i veštačke inteligencije nije ni malo ohrabrujuća jer umesto da obezbedi veću slobodu i ispunjenje od onoga što je doneo dosadašnji kapitalizam, tehnološki napredak zasnovan

na AI plaši nas sa još većim oblikom beznađa, strahom od beskorisnosti. WEF nudi alternative u takozvanom “inkluzivnom kapitalizmu” u kome će AI imati emancipatorski karakter kroz instituciju univerzalnog dohotka.

Istorijat ideje o univerzalnom osnovnom dohotku (Univezal Basic Income – UBI) indirektno potiče još od ekonomiste Džona Majnarda Kejnsa i ideje da (svetska) vlada treba da kontroliše sve aspekte ekonomije po svom izboru kako bi sistem nastavio da funkcioniše, pri čemu sloboda pojedinca nije u fokusu. Kejns je verovao da je uloga vlade da osmisli „promišljenu nacionalnu politiku o tome koja je veličina stanovništva najcelishodnija. I kada usvoji ovu politiku, vlada mora preduzeti korake da je sprovede u delo. Možda će doći vreme kada zajednica u celini mora da obrati pažnju na urođeni kvalitet kao i na puki broj svojih budućih članova” (Ammous, 2018). Sve ovo podseća na osnovne postavke eugenike, koja se danas veoma često pojavljuje u javnom diskursu u kontekstu mogućih ishoda „četvrtе industrijske revolucije” i “Velikog reseta”, koji se uveliko oslanjaju na AI. Ovakva situacija predstavlja veliki bezbednosni rizik na globalnom svetskom nivou, pretnju od masovnih socijalnih nemira, jer se ukida osnovno ljudsko pravo na rad koji obezbeđuje dostojanstvenu egzistenciju.

Veza AI i UBI, koja bi mogla da spereči ove rizike, ogleda se u novim ekonomskim modelima zasnovanim na DLT (Distributed Ledger Tecnologies). DAO (Decentralized Autonomous Organization) je nova organizaciona forma na blokčejnu, računarski algoritam u formi pametnog ugovora (smart contract) koji primenjuje prava vlasništva nad kapitalom, ugovorne obaveze i pravila poslovne logike. Vlasnici kripto valute akumuliraju moć i kapital osnivanjem organizacija sa sopstvenim novcem i na taj način stiču moć odlučivanja. Danas kada je na delu ANI (Artificial Narrow Intelligence) kodove ispisuju ljudski programeri – koderi. Idealno rešenje je ono u kome bi veštačka inteligencija mogla da traži različite parametre i donosi najbolje moguće odluke, da kreira sopstvene proizvode i usluge i da ih prodaje, dok bi profit išao ljudskim bićima, onim „nezapošljivim” kako ne bi stekli status „beskorisnih”. Tako u budućnosti, AI DAO mogu igrati veoma značajnu ulogu u konceptu

implementacije UBI. Za sada je to nemoguće postići jer još uvek nije stvorena veštačka opšta inteligencija (AGI) koja je sposobna da razume svet kao i svaki čovek, i sa istim kapacitetom da nauči kako da izvrši ogroman broj zadataka u realnom vremenu i kompleksnom okruženju. Danas je veštačka inteligencija na radnom mestu zastupljena kroz sisteme za praćenje, sa novim prediktivnim kapacitetima u cilju uspostavljanja sve invazivnijih mehanizama upravljanja radnicima, kontrole imovine i ekstrakcije vrednosti.

AI i resursi

Sistemi mašinskog učenja zahtevaju ogromne količine podataka. Na primeru interpretacije slika radi prepoznavanja lica i tumečenja emocija videćemo koliko je duboko složen i relacioni taj poduhvat. Slike su izuzetno osetljivi podaci, opterećene višestrukim potencijalnim značenjima, tumačenjima i kontradikcijama. Danas je uobičajena praksa da se kao prvi koraci u kreiranju sistema kompjuterskog vida prikupljaju milioni slika sa interneta, poređaju u niz klasifikacija i koriste kao osnova za način na koji će sistem sagledati vidljivu stvarnost. Ove ogromne kolekcije se nazivaju skupovi podataka za obuku ili žargonski „osnovna istina“ (Jaton, 2017). Istina se, dakle, oslanja na gomilu slika prikupljenih iz raznih izvora na mreži. Skupovi podataka za obuku su, dakle, srž načina na koji većina sistema mašinskog učenja donosi zaključke. Oni služe kao primarni izvorni materijal koji sistemi veštačke inteligencije koriste da formiraju osnovu svojih predviđanja. Postoji više modela mašinskog učenja, kao na primer neuronske mreže, stabla odlučivanja, logistička regresija itd. “Za nadgledano mašinsko učenje, inženjeri postavljaju prikupljene podatke o obuci na računar. Dva različita tipa algoritama tada stupaju u igru: učenici i klasifikatori. Učenik je algoritam koji je obučen ovim inicijalnim skupom podataka; on “obaveštava” klasifikatora kako najbolje analizirati odnos između novih ulaza” (Krawford, 2021), odnosno podataka prikupljenih nadzorom, i željenog ciljanog izlaza (ili predviđanja). Što više primera ispravno označenih podataka ima, algoritam bi trebao biti bolji u stvaranju tačnih predviđanja, odnosno u ovom slučaju prepoznavanju lica

ili emocija.

Jednom kada se setovi za obuku postave kao korisna merila, oni se obično prilagođavaju, nadograđuju i proširuju. Podaci o obuci su, dakle, osnova na kojoj su izgrađeni savremeni sistemi mašinskog učenja. “Ali podaci o obuci su krhki oblik temeljne istine — pa čak ni najveće baze podataka ne mogu izbeći fundamentalne greške koje nastaju kada se beskonačno složen svet pojednostavi i klasifikuje u kategorije” (Krawford, 2021). U suštini, svako mašinsko učenje se odnosi na svođenje sa konteksta na podatke, sa značenja na statističko prepoznavanje obrazaca.

Kancelarija za razvoj tehnologije za borbu protiv droga Ministarstva odbrane SAD sponzorisa je program tehnologije prepoznavanja lica Facial Recognition Technology (FERET) za razvoj automatskog prepoznavanja lica za obaveštajne službe i sprovođenje zakona. Uz saglasnost učesnika i u kontrolisanim uslovima, snimljeni su portreti više od hiljadu ljudi, u više poza, kako bi se napravilo ukupno 14.126 slika u cilju „otkrivanja prevara“ kroz višestruke zahteve za socijalnu pomoć, kao i za kontrolu granica i aerodroma radi identifikovanja već poznatih krijumčara, terorista ili drugih kriminalaca. To je bilo pre nego što je internet počeo da “nudi” masovno i besplatno vađenje slika bez ikakvih dozvola i saglasnosti, tretirajući podatke slično prirodnom resursu – slobodno za eksploataciju. Odjednom su setovi za obuku mogli da dostignu veličinu koju naučnici 1980-ih nisu mogli ni da zamisle. “Nestala je potreba za snimanjem fotografija jer sada su postojali milioni selfija u svim mogućim uslovima osvetljenja, položaja i dubine polja. Ljudi su počeli da dele fotografije svojih beba, porodične snimke i slike kako su izgledali pre deceniju, što je idealan izvor za praćenje genetske sličnosti i starenja lica” (Krawford, 2021), kao jedan od parametara za pisanje savršenijih algoritama za prepoznavanje lica i emocija. Na primer, jednog prosečnog dana u 2019. godini, na Fejsbuk je postavljeno oko 350 miliona fotografija (Krawford, 2021). Kada su slike skinute sa interneta, pojavila se velika briga: ko bi ih sve označio i stavio u razumljive kategorije od kojih će se formirati skupovi podataka za obuku. U tu svrhu angažovane su armije slabo plaćenih

radnika da sortiraju u proseku pedeset slika u minutu u hiljade kategorija. Pristup masovnog vađenja podataka bez saglasnosti i klasifikacije od strane nedovoljno plaćenih i nestručnih radnika postao je standardna praksa u mnogim kompanijama koje su kreirale i zatim prodavale skupove podataka za obuku AI. Iz svakog naizgled bezazlenog i anonimnog skupa podataka može se doći do mnogih neočekivanih i veoma ličnih informacija, što nije omelo još intenzivnije i bezobzirnije prikupljanje slika i drugih podataka sa interneta. Podaci su u dvadeset prvom veku postali ono što je nafta značila u dvadesetom veku. "Podaci su počeli da se opisuju kao resurs koji treba potrošiti, tok koji treba kontrolisati ili investicija koju treba iskoristiti" (Krawford, 2021). Internet i WEB2.0 direktno su odgovorni za enormnu koncentraciju moći nekoliko platformi kao što su Facebook, Amazon, YouTube, Twitter i Google. Ove platforme stekle su ne samo medijsku, ekonomsku već i ogromnu političku moć omogućivši novo doba nadzornog kapitalizma i "online svet koji ne podleže zemaljskim zakonima" (Zubof, 2020). "Google kao pionir, izumitelj nadzornog kapitalizma" (Zubof, 2020) prisiljava svoje korisnike da trampe svoje bihevioralne podatke za uslugu informisanja i povezivanja. Ti podaci se metodom mašinskog učenja ili veštačke inteligencije pretvaraju u visokoprofitabilne algoritamske proizvode kojima se predviđa ponašanje Guglovih korisnika. Jedan deo tih podataka predstavlja "bihevioralni višak koji postaje temelj sasvim nove logike akumulacije" (Zubof, 2020), ne samo u komercijalne svrhe već i kao glavno oruđe za društveni nadzor. Nadzorni kapitalizam postao je osnovni model informacionog kapitalizma na mreži.

Pristrasni skupovi podataka kreiraju nepravedni algoritam-klasifikator čija primena u praksi prepoznavanja lica ili emocija može da ima neželjene posledice koje štete ranjivim zajednicama i jačaju već ukorenjene nepravde i diskriminacije. Reprodukacija štetnih algoritama je posebno opasna sada kada je veštačka inteligencija od eksperimentalne discipline koja se koristila samo u naučnim laboratorijama postala testirana i primenjena na milionima ljudi u realnom svetu. Prakse ekstrakcije podataka zalaze u oblasti ljudskog života koje su nekada bile zabranjene ili nepristupačne. Tehnološke kompanije koriste

nove pristupe kako bi došle do glasovnih podataka preko kućnih uređaja, podatke o našem kretanju preko satova na ruci i telefona u džepovima, sa tableta i laptopa uzimaju podatke o našim interesovanjima i preferencijama, do gestova i mimika dolaze preko kamera na radnim mestima i učionicama itd. Ti podatci neretko se ustupaju sumnjivim vladinim agencijama na osnovu javno-privatnih partnerstava što ukazuje da je prikupljanje i označavanje podataka o svetu društvena i politička intervencija, čak i kada se prikazuje kao čisto tehnički postupak. Celokupna praksa prikupljanja podataka, njihovog kategorisanja i označavanja, a zatim korišćenja za obuku sistema indirektno je oblik politike.

AI i prepoznavanje emocija

Postupak klasifikacije podataka je osnovna praksa u veštačkoj inteligenciji, ali “ono što često nedostaje je fundamentalniji skup pitanja: kako klasifikacija funkcioniše u mašinskom učenju? Na koje načine klasifikacije stupaju u interakciju sa klasifikovanim? Koje neizrečene društvene i političke teorije leže u osnovi i koje podržavaju ove klasifikacije sveta” (Krawford, 2021). Klasifikacija je čin moći, međutim industrija veštačke inteligencije tradicionalno shvata problem pristrasnosti prilikom klasifikacije kao da je to sporadična greška koju treba popravljati od slučaja do slučaja, a ne karakteristika same klasifikacije. Na primer, IBM-ovi istraživači svoje klasifikacije zasnivaju na krajnje problematičnim i nenaučnim pretpostavkama, kao što je pretpostavka “da se aspekti našeg nasleđa – uključujući rasu, etničku pripadnost, kulturu, geografiju – i naš individualni identitet – godine, pol i vidljivi oblici samoizražavanja – odražavaju na našim licima” (IBM T. J. Watson Research Center, 2019). Ovakvi sistemi klasifikacije mogu svojom primenom u sistemima zasnovanim na AI naneti ogromnu štetu ljudima i čitavim zajednicama. Ovakva redukcionističko-biologistička shvatanja “delikventnu osobu vide kao ličnost: specifičnih telesnih degenerativnih karakteristika, bilo konstitucionalnih posebnosti, endokrinih i hromozomskih aberacija, genetskih predispozicija, urođenih sklonosti ili rasnih predispozicija” (Bošković, 2012). Na primer, kategoriju

“kriminalnost” koja je relaciona i društveno uslovljena, AI sistemi za prepoznavanje nastoje da klasifikuju u fiksne kategorije kako bi postigli “pouzdana” izlaze pretrage za potrebe policijskih službi. Ovakvi pristupi s pravom se kritikuju kao naučno i etički problematični.

Ideja o tehnikama prepoznavanja emocija polazi od pretpostavke o mogućnosti anatomsko-fiziološkog razumevanja lica preko “mikroekspresije - sitnih pokreta mišića lica” (Ekman and Friesen, 1969). Istraživač Pol Ekman je kodifikovao sistem za otkrivanje i analizu izraza lica i objavio ga pod imenom Facial Action Scoring Technique (FAST), Tehnika ocenjivanja facijalnih akcija. Ekman je na osnovu poziranih fotografija u kontrolisanim uslovima identifikovao otprilike četrdeset različitih mišićnih kontrakcija na licu i nazvao osnovne komponente svakog izraza lica “akcionom jedinicom” (Ekman and Friesen, 1969). Iz njih je izveo šest osnovnih emocionalnih stanja. Nakon testiranja objavio je Facial Action Coding System (FACS), Sistem za kodiranje facijalne akcije, čija se ažurirana izdanja dopunjena podacima iz spontanih izraza lica koriste i danas (Ekman and Rosenberg, 2005). Čak i kada su slike sakupljene u prirodnim okruženjima, one se obično klasifikuju prema Ekmanovoj šemi izvedenoj iz poziranih slika. Ekman je postao legenda “čitanja lica”, tako da je Američkoj upravi za bezbednost saobraćaja izradio SPOT program (Screening of Passengers by Observation Techniques), korišćen za praćenje izraza lica putnika avionom u godinama nakon napada 11. septembra, pokušavajući da „automatski” otkrije teroriste. Program SPOT uprkos ceni od devetsto miliona dolara, nije doneo jasne rezultate.

Kejt Krošel, menadžer za automatizaciju marketinga u kompaniji iMotions, objašnjava: „Najčešće referencirana metodologija je analiza ekspresije lica, koja radi na algoritmima obučanim od strane koderica ljudskih lica i milionima ljudskih lica i njihovim akcijama kao što su uvijanje usana, podizanje obrva, brazdi i bora, itd. Tehnologija AI spaja ove akcione jedinice kako bi procenila ekspresivnost sedam osnovnih emocija (radost, bes, strah, gađenje, prezir, tuga i iznenađenje) i verovatnoću da se očekivana emocija izražava.” Današnja

tehnologija za prepoznavanje emocija je vrsta veštačke inteligencije koja se odnosi na prepoznavanje lica na osnovu ovih sedam navedenih osnovnih emocija, i koja pokušava da identifikuje emocionalna stanja čoveka na osnovu njegovih izraza lica i telesnih znakova, uključujući puls i aktivnost mozga. Softver takođe može da prati kretanje očiju kako bi identifikovao na koje delove stimulusa subjekt obraća najveću pažnju.

Upotreba AI za prepoznavanje emocija u cilju prevencije nasilnog ponašanja i terorizma koristi se i danas uprkos brojnim kritikama o naučnoj neutemeljenosti i narušavanju građanskih sloboda. Postoje i neke druge nedoslednosti po pitanju same utemeljenosti AI tehnika za prepoznavanje emocija jer među istraživačima čak ne postoji ni konsenzus o tome šta su zapravo emocije, kako su u nama formulisane i izražene, koje bi mogle biti njihove fiziološke ili neurobiološke funkcije, njihov odnos prema stimulansima itd. „Reč je o tome da su emocije najčešće izazvane vrednovanjem ili procenom događaja, u vezi sa svim onim što smatramo važnim: našim ciljevima, našim brigama i težnjama” (Outli, 2005). Termin emocija obuhvata širok spektar fenomena pa se koriste i izrazi afekat, sentiment, preferencija. „One daju intenzitet našem životu i jednim su delom zasnovane na našim kulturnim kontekstima i onome što smatramo društveno prihvatljivim” (Outli, 2005). Naše ponašanje, pa i izrazi lica koji ih prate, vrlo retko su odraz čisto emocionalne reakcije, već su odraz kombinacije racionalne procene situacije i emocionalnog odgovora na nju. Zbog toga se one manifestuju u bezbroj nijansi, intenziteta i formi koje se očituju kao naše facijalne ekspresije. To ih često čini neuhvatljivim i za iskusno oko obučenog posmatrača, a kamoli za svedeni algoritam AI sistema za nadzor i procenu emocija, pa se postavlja pitanje da li se emocije uopšte mogu adekvatno grupisati u mali broj diskretnih kategorija.

Istraživač afektivnog računarstva Arvid Kappas, navodi: „Premalo znamo o složenim društvenim modulatorima facijalne i eventualno druge ekspresivne aktivnosti da bismo mogli pouzdano da izmerimo emocionalno stanje na osnovu ekspresivnog ponašanja. Ovo nije inženjerski problem koji bi se mogao rešiti boljim algoritmom” (Kappas, 2010). Uprkos tome, mnogi skupovi

podataka za obuku AI sistema zasnivaju se na glumcima koji simuliraju emocionalna stanja, što znači da su AI sistemi obučeni da prepoznaju lažne odnosno neiskrene izraze osećanja. Iako sistemi veštačke inteligencije tvrde da imaju pristup temeljnoj istini o prirodnim unutrašnjim stanjima, oni su obučeni na materijalu koji je evidentno konstruisan.

Liza Feldman Barret je na osnovu opsežnog pregleda literature o zaključivanju emocija na osnovu izraza lica konstatovala da „nije moguće sa sigurnošću zaključiti sreću iz osmeha, ljutnju ili tugu iz mrštenja, kao što većina trenutne tehnologije pokušava da uradi kada primenjuje ono za šta se pogrešno veruje da predstavlja naučne činjenice“ (Barret, 2006). To znači da su ekspresije mešavina različitih emocija koje im daju bezbroj nijansi i opiru se redukcionističkim klasifikacijama raznih „naučnih istraživanja“. Prepoznavanje emocija putem AI sistema može biti u potpunosti pogrešno ako je bazirano na pretpostavci da su emocionalne kategorije date, a ne pojavne i relacijske. „Prerano je koristiti ovu tehnologiju za donošenje zaključaka o tome šta ljudi osećaju na osnovu pokreta lica“ (Barret, 2006). Uprkos ovoj činjenici koju priznaje sve više naučnika, nova generacija AI alata za afekte se primenjuju u sve većem rasponu konteksta, od rada policije u otkrivanju i prevenciji nasilničkog ponašanja, do selekcije kandidata prilikom zapošljavanja, odnosno šire se onoliko brzo koliko laboratorije i korporacije mogu da stvore nova tržišta za njih.

Nasilničko ponašanje i tumačenje emocija putem AI

Činjenica je da otkada postoji čovek postoji i nasilje. U vremenima kad je čovekova svest bila usmerena na borbu za opstanak, agresivno ponašanje je bilo korisno u odbrani i napadu jer su preživljavali oni borbeniji i jači. Ponekad je nasilje bilo posledica straha od različitosti, nerazumevanja tuđeg pogleda na svet ili kulture. Na globalnom planu, tokom cele istorije čovečanstva, gotovo da nije bilo perioda kada se nije vodio neki rat i kada nasiljem nije nastojala da se stekne neka korist. Zaista, uvek je bilo nasilja proizašlog iz pogrešnih uverenja, osećaja ugroženosti, pohlepe za tuđim... No, ono što

naše vreme čini drugačijim i što zastrašuje jeste eskalacija nasilja u kome svako može postati žrtva bezumnog, besmislenog nasilja koje je samo sebi postala svrha (Bjelajac & Matijašević, 2013). Posledice koje izaziva, ne utiču samo na pojedinca i njegov lični razvoj već imaju višestruke negativne implikacije po užu i širu socijalnu sredinu i društvo u celini (Merdović & Bjelajac, 2021). Nažalost, najčešće žrtve su žene i deca, (Bjelajac & Merdović, 2019), koje su sada naročito izložene sveopštem digitalnom nasilju, koje fluktuiraju između virtuelnog i stvarnog (Bjelajac & Filipović, 2021).

Još u dvanaestom veku, pre pojave psihologije i ostalih nauka koje se bave ljudskim ponašanjem i emocijama, monarh Evagrius napravio je katalog od sedam smrtnih grehova koji okupiraju ljudsku dušu. Tu spadaju pohlepa, ponos, proždrljivost, požuda, bes, lenjost i zavist. Pohlepa je dominantno obeležje duše i iz nje proizilaze ostali grehovi. Ovakva „stanja duše“ prate emocionalni odgovori kao i korespondentni facijalni izrazi. Slično pohlepi, mnogi naučnici tvrde da je strah bazična urođena emocija kod čoveka, i on proističe iz potrebe da se zaštiti telesni integritet i život sa svim svojim aspektima koje osoba smatra za važne. Iz straha proizilaze sve ostale negativne emocije, dok je odsustvo straha preduslov za ispoljavanje pozitivnih emocija.

Pored različitih tumačenja osnovnih emocionalnih stanja koja imaju specifičan facijalni izraz, prihvatanja postojanja i drugih afekata, činjenice da su osećanja i emocije najčešće „pomešani“ i kompleksni izrazi reagovanja na situaciju, postavlja se jedno fundamentalnije pitanje: da li je moguće na osnovu emocionalnih izraza lica i ostalih telesnih manifestacija predviđati ponašanje ljudi, karakterisati ih putem AI i konačno preduzimati akcije (bilo preventivne bilo korektivne) protiv njih?

Profil nasilničke ličnosti je određen kako socijalno-relacionim tako i biološko-psihološkim karakteristikama. Nasilje ima fizičko-telesne manifestacije i emocionalna stanja u korespondentnim facijalnim izrazima. „Osoba sklona nasrtljivosti tipologizovana je u teoriji kao agresivna ličnost. To je netolerantana, nasrtljiva osoba, sklona konfliktima u rešavanju sporova nasiljem. Manifestuje se kao antisocijalni poremećaj neurotičnih svojstava sa jakim osećajem

sopstvene nesigurnosti, dubokog nepoverenja u druge i neprestanog straha od vlastite ugroženosti u društvenoj sredini” (Bošković, 2010). U ovoj definiciji nasilničke agresivne ličnosti vidimo njene pojavne fizičke manifestacije ali i emocionalna stanja koja ih prate, kao što su nesigurnost, nepoverenje, strah. Iako agresivnost može biti i pozitivno, pa čak i poželjno svojstvo u nekim situacijama, ono najčešće ima negativnu konotaciju u ljudskom ponašanju.

Ako pođemo od Ekmanovih negativnih bazičnih emocija gde spadaju bes, strah, gađenje, prezir, tuga, vidimo da je iz karakterizacije agresivne ličnosti prisutan jedino strah, dok nesigurnost i nepoverenje možemo tumačiti kao „ostali afekti” koji proizilaze iz straha kao preovlađujuće emocije. Međutim, strah je univerzalna emocija i nikako nije povezana jedino sa agresivnošću, kao što na primer ni olakšanje i zadovoljstvo nisu povezani sa tolerantnim i prijateljskim ponašanjem. Na primer, nasilnik prilikom vršenja dela uopšte ne mora da ispoljava strah, a nakon počinjenog dela može da oseća olakšanje i zadovoljstvo, dok naprotiv, žrtva izražava ceo dijapazon negativnih emocija kao što su strah, bes, prezir, tugu i sl. „Od svih oblika poremećaja ličnosti za delikte nasilja najveći značaj imaju psihopatije. Kod psihopatije nema poremećaja inteligencije ni sposobnosti rasuđivanja, već je reč o izopačenosti u moralnim nazorima, poimanju socijalnih normi, a posebno emotivnih komponenti”. To znači da pri činjenju nasilničkih dela nemamo nikakvu ili odgovarajuću emociju, te emocionalna stanja i njihni manifestni facijalni izrazi ne mogu biti potpuno pouzdana osnova za tumačenje nasilničkog mentaliteta, njihovu karakterizaciju, kao osnovu za preduzimanje akcija prema njima. To se pogotovo odnosi na upotrebu AI sistema koji su, bez obzira na bogatstvo i veličinu skupova podataka za obuku i savršenosti algoritma, uvek redukovani i pojednostavljeni. Upotreba AI sistema za prevenciju nasilničkog ponašanja pored ostalih elemenata kao što su baze podataka već evidentiranih nasilnika, geografskih lokacija učestalosti nasilja, susedskih relacija i još mnogo potencijalnih baza podataka (društvene mreže, google mape, GPS praćenja preko telefona...), treba da sadržati i što je moguće preciznije konstruisane algoritme za prepoznavanje

emocija, samo kao jedan od faktora pri proceni rizika od nasilničkog ponašanja.

Zaključak

Iz svega prethodnog možemo zaključiti da je na osnovu dosadašnjeg, javnosti dostupnog saznanja, domet AI relativno skroman ali da s obzirom na neverovatno brz tehnološki razvoj, naročito u oblasti kvantnog računarstva, pred ovom oblasti stoji veoma dinamična budućnost. Kvantno računarstvo je osnova za dalji razvoj AI, s obzirom na količinu podataka koje može da analizira jer je zasnovano na "superpoziciji" elementarnih čestica koje se kreću oko jezgra atoma, što znači da se informacija zapisuje u bezbroj mogućih stanja atoma (kubit) a ne samo kao binarni kod, jedinica i nula (bit) koje korespondiraju uključenom-isključenom strujnom kolu. To znači daleko savršenije algoritme, bogatije i raznovrsnije baze podataka (*big data*), *deep learning*, što će za krajnji rezultat imati daleko pouzdanije izlaze nego što su to danas. Imaće svakako i duboke ekonomske, političke, bezbednosne, psihološke i socijalne reperkusije. Sve ove nabrojane tehnologije leže u osnovi Velikog reseta, prve zaista sveobuhvatne globalne transformacije sveta kakvog smo do sada poznavali. Da li je to put u bolju budućnost ili samo jedna avantura čovečanstva koje ubrzano klizi ka neizvesnoj distopiji društva apsolutnog nadzora, vreme će veoma brzo pokazati.

Reference

- Ammous, S. (2018). *The Bitcoin standard* [Bitcoin standard]. John Wiley & Sons, Inc.
- Barret, F.L. (2006). *Are Emotions Natural Kinds? Perspectives on Psychological science* [Da li su emocije prirodni načini? Pogledi na psihologiju]. Boston College.
<https://www.affective-science.org/pubs/2006/Barrett2006kinds.pdf>
- Bjelajac, Ž., & Matijašević, J. (2013). Nasilje kao oblik socio-patološkog ponašanja. *Kultura polisa*, 10(22), 409-425.
<https://kpolisa.com/index.php/kp/article/view/1278>
- Bjelajac, Ž., & Merdović, B. (2019). Nasilje nad decom u porodici. *Kultura polisa*, 16(39), 191-202;
<https://kpolisa.com/index.php/kp/article/view/343>
- Bjelajac, Ž. & Filipović, A. (2021), Specific characteristics of digital violence and digital crime [Specifičnosti digitalnog nasilja i digitalnog kriminala]. *Pravo – teorija i praksa*, 38(4), 16-32.
<https://doi.org/10.5937/ptp2104016B>
- Bošković, M. (2012). Psihosocijalne karakteristike nasilničke ličnosti. *Zbornik radova Pravnog fakulteta u Novom Sadu*. XLVI(1), 7-21.
<http://zbornik.pf.uns.ac.rs/images/download/2012/2012-1.pdf>
- Crawford, K. (2021). *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence* [Atlas AI. Moć, politika i planetarni troškovi veštačke inteligencije]. Yale University press.
- Dreyfus, L. H. (1999). *What computers still can't do* [Šta računari još uvek ne mogu da urade]. The MIT press.
- Ekman, P., & Friesen, W. V. (1969). Nonverbal Leakage and Clues to Deception [Neverbalno curenje i tragovi za prevaru]. Reprinted from *Psychiatry journal for the study of interpersonal processes*.

<https://www.paulekman.com/wp-content/uploads/2013/07/Nonverbal-Leakage-And-Clues-To-Deception.pdf>

Ekman, P. & Rosenberg W.E, (2005). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)* [Šta lice otkriva: osnovne i primenjene studije spontanog izražavanja pomoću sistema kodiranja facijalnih akcija (FACS)]. Oxford University Press.

<https://www.aqualide.com/upload/texte/text98.pdf>

Merler, M., Ratha, N., Feris, R., & Smith, J. R. (2019, April 10). Diversity in Faces [Raznolikost u licima]. *arXiv*: 1901.10436.

<https://arxiv.org/pdf/1901.10436.pdf>

Jaton, F. (2017). We get the algorithms of our ground truths: Designing referential databases in digital image processing [Dobijamo algoritme naših osnovnih istina: Dizajniranje referentnih baza podataka u digitalnoj obradi slike]. *Social Studies of Science*, 47(6) 811–840.

<https://doi.org/10.1177/0306312717730428>

Kappas, A. (2010). Smile When You Read This, Whether You Like It or Not: Conceptual Challenges to Affect Detection [Nasmejte se kada ovo čitate, sviđalo vam se to ili ne: konceptualni izazovi koji utiču na otkrivanje]. *IEEE Transaction on affective computing*, 1(1) 38–41.

https://www.academia.edu/1180128/Smile_when_you_read_this_w_hether_you_like_it_or_not_Conceptual_challenges_to_affect_detectio
<https://doi.org/10.1177/0306312717730428>

Merdović, B., & Bjelajac, Ž. (2021). Multidimenzionalne perspektive nasilja u porodici u Srbiji. *Kultura polisa*, 18(46), 185–202.

<https://doi.org/10.51738/Kpolisa2021.18.3r.3.01>

Outli, K. (2005). *Emocije*. Clio.

Zubof, Š. (2020). *Doba nadzornog kapitalizma*. Clio.

Artificial Intelligence (AI) in the Function of Recognizing Emotions and Violent Mentality

Momčilo B. Bajac¹ and Željko Đ. Bjelajac²

¹University UNION Nikola Tesla in Belgrade
Faculty of management, Sremski Karlovci

²University Business Academy in Novi Sad
Faculty of Law for Commerce and Judiciary in Novi Sad

Summary

In this paper, we discuss the impact of the latest technologies, primarily artificial intelligence (AI), on the wider social and natural environment, as well as on the individual lives of people in an increasing range of contexts. We talk about AI as complex ecosystems made up of natural resources, energy, human labour, infrastructure, logistics, history, data, and a multitude of classifications. Even though for now, AI is neither artificial nor as intelligent as it is presented to the general public, there is a strong desire for its universalization and absolutization. In this paper, we have specifically discussed the possibility of applying AI to recognize emotions and violent mentality to prevent violent behaviour. This required an analysis of the practice of collecting data and classifying them into fixed categories in order to achieve “reliable” search outputs for different needs. The profile of a violent personality is determined by both social-relational and biological-psychological characteristics, which represents a large number of variables that should be taken into account when constructing an algorithm for recognizing emotions and violent personality. We conclude that with the current level of AI

development and very reduced and discrete algorithms, it is not possible to take into account enough data to obtain a reliable way to predict behaviour and take possible measures without discrimination and danger to individuals and society as a whole. But given the incredibly rapid technological development, especially in the field of quantum computing, this field faces a very dynamic future.

Keywords: algorithm, datasets, classification, facial expressions, violent personality, quantum computing