

# (Ne)Razuman strah od veštačke inteligencije

---

Damjan Krstajić

Pre četiri godine grupa istaknutih naučnika i preduzetnika, među kojima su bili Stivn Hoking (Stephen Hawking) i Elon Musk, je potpisala otvoreno pismo o veštačkoj inteligenciji u kojem traže da se istraži uticaj ove tehnologije na društvo. Iako se u pismu, koji je objavio The Future of Life Institute, spominje i dobrobit od veštačke inteligencije, glavna vest su ipak bila ozbiljna upozorenja za budućnost. Mišljenja sam da već danas imamo dovoljno materijala za to.

Prošle nedelje sam probao da demistifikujem veštačku inteligenciju i da objasnim da postoji naučna oblast pod tim nazivom, ali ne i sama veštačka inteligencija. Ono što nam se danas većinom nudi kao proizvod veštačke inteligencije je rezultat mašinskog učenje, a to je opet neka vrsta statističkog uopštavanja nastala obradom ogromne količine podataka. Tu nema ni traga kognitivnim procesima. Po meni, primereniji naziv za sve te alate je *pomoćna inteligencija*. Kako je ona bazirana na podacima, onda ovde važe i sva upozorenja vezana za zaključivanje na osnovu statistike. Ako su kreatori i korisnici tih alata svesni toga, onda možemo da govorimo o dobrobitima.

Pomoćna inteligencija u personalizovanoj medicini (moja oblast istraživanja) je primer pozitivne primene novih tehnologija gde su proizvodi dodatni alat lekarima prilikom donošenja odluka. Evo primera. Metastatski tumori nepoznatog porekla su ozbiljan problem u onkologiji, jer su to situacije kad imamo pronađen karcinom, ali ne i odakle je metastazirao. Pre petnaestak godina grupa naučnika je oformila firmu u SAD sa ciljem da uz pomoć nove tehnologije genetskih ekspresija predvidi organ odakle je došao tumor. Problem je što eksperimentalni rezultati ovih genetskih ekspresija izlaze u ljudski nečitljivom formatu od preko 20 000 brojeva. Ljubomir Buturović, doktor nauka sa beogradskog univerziteta koji živi i radi u SAD, je bio na čelu tima za pravljenje predikcionog modela. Dobrim dizajnom eksperimenta i uz pomoć mašinskog učenja uspeali su da kreiraju model koji od genetske ekspresije kvalitetno predviđa organ porekla tumora. Uspeh se

ovde ne meri u objavljenim radovima, već to što je model prošao rigorozne testove američke agencije FDA i danas je u kliničkoj upotrebi.

Ozbiljan problem nastaje kad se pomoćna inteligencija ne koristi za dobrobit čovečanstva i kada njihovi kreatori ili korisnici nisu svesni upozorenja o neadekvatnoj upotrebi statistike. Po meni, osnovni problem je što se slabo razume pristrasnost koja postoji u opservacionim podacima.

Da li biste voleli da odluku o visini kazne u sudu odlučuje predikcioni algoritam (takozvana veštačka inteligencija) ili čovek? U sudovima u SAD se već nekoliko godina koristi softver COMPAS koji na osnovu 137 podataka o osuđeniku određuje rizik od ponovnog prekršaja, a prošle godine britanska vlada je poslala predlog za primenu automatskog sistema kažnjavanja za sitne kriminalne radnje. Uz pomoć informacija o prethodnim prekršiocima zakona i na osnovu podataka o novom osuđeniku, izračunava se rizik od ponovnog prekršaja i na osnovu toga određuje dužina kazne.

Da li biste voleli da kad pošaljete svoju biografiju firmi, taj dokument prolazi kroz sistem takozvane veštačke inteligencije koji odlučuje da li ćete biti pozvani na intervju? Na osnovu svih do sada pristiglih biografija i uzimajući u obzir one kandidate koje su oni na kraju zaposlili, američka kompanija Amazon je napravila i primenila model za predviđanje uspešnih kandidata i tako uštedila pare.

Oba ova primera su se kasnije pokazala pristrasnim, iako njihovim kreatorima, pretpostavljam, to nije bio cilj. U prvom slučaju se ispostavilo da crnci dobijaju strožije kazne, a u drugom da žene bivaju ređe pozivane na intervju. Problem je u zbunjujućim faktorima (confounding). Crnci su u zatvorima u SAD proporcionalno više zastupljeni nego u populaciji. Muškarci su bili proporcionalno više zastupljeni među zaposlenima u Amazonu nego u populaciji. Možete iz vaših podataka da izbacite informacije o rasi i polu, ali to ne znači da oni i dalje nemaju uticaja u analizi. Ako je automatski sistem za određivanje kazni našao kalup da osuđenici iz kraja A imaju veće šanse za ponovno nedelo od osuđenika iz kraja B, onda će stanovnicima kraja A dati strože kazne. Ako je sistem za filtriranje kandidata za intervju našao kalup u podacima da su kandidati koji u slobodno vreme upražnjavaju američki fudbal bili više zapošljavani od onih sa drugim hobijima, onda će ovima koji igraju američki fudbal dati veće šanse za poziv. Kako sistem može da zna da u kraju A žive većinom crnci, ili da malo žena igra američki fudbal?

Dugo mi nije bilo jasna ideja o autonomnim vozilima. Sistemi za prepoznavanje su bazirani na mašinskom učenju i na autonomno vozilo

možemo da gledamo kao na sistem za odlučivanje koji koristi puno različitih statističkih modela. Koliko god da su oni sjajni uvek će postojati neki slučaj koji nije do tada poznat i samim tim mogućnost pogrešne odluke i fatalnog ishoda. Evo primera. Vozite se autoputem i u kolima ispred vidite decu pozadi kako se igraju i izvlače STOP znak. Deca se igraju, šala, nastavljate dalje, dok autonomno vozilo vidi STOP i naglo zaustavlja u sred autoputa. Neko će reći, u redu, ispravićemo softver da prepozna STOP u pokretnom objektu i da onda ne stane naglo. Dobro, ali koliko ima ovakvih mogućnosti i kako sve njih isprogramirati?

Zahvalan sam kolegi koji me je upozorio na naivnost po ovom pitanju. Šta mislite koliko truda treba da se promeni softver autonomnog vozila da bi se on koristio u vojne svrhe? Umesto instrukcije „vidiš pešaka ispred, zaustavi se“ da bude „vidiš pešaka ispred, nastavi dalje (gazi)“. Kolega tvrdi par linija softver koda je dovoljna za ovu promenu.

Ne plašim se veštačke inteligencije, jer ona u kognitivnom smislu ne postoji, već ljudi koji u njeno ime koriste softver da iz bilo kog razloga skinu odgovornost za odluke o tuđim životima. Plašim se svoje i tuđe naivnosti po pitanju objektivnosti onog što nam se nudi kao veštačka inteligencija.

Reference koje podržavaju činjenice spomenute u članku

1. AI Open Letter

<https://futureoflife.org/ai-open-letter/>

[https://en.wikipedia.org/wiki/Open\\_Letter\\_on\\_Artificial\\_Intelligence](https://en.wikipedia.org/wiki/Open_Letter_on_Artificial_Intelligence)

2. Personalizovana medicina

[https://en.wikipedia.org/wiki/Personalized\\_medicine](https://en.wikipedia.org/wiki/Personalized_medicine)

3. Metastazirani tumori nepoznatog porekla

<https://www.cancerresearchuk.org/about-cancer/cancer-unknown-primary-cup>

4. Ljubomir Buturović

<https://www.linkedin.com/in/ljubomir-buturovic-798156>

5. Tissue of Origin Test

<http://www.cancergenetics.com/laboratory-services/specialty-tests/too-tissue-of-origin-test/>

6. U sudovima u SAD se koristi COMPAS softver

[https://en.wikipedia.org/wiki/COMPAS\\_\(software\)](https://en.wikipedia.org/wiki/COMPAS_(software))

7. Više o COMPAS softveru

<http://advances.sciencemag.org/content/4/1/eaao5580>

8. O tome kako SAD planiraju da koriste veštačku inteligenciju u sudovima

<https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html>

9. Planovi britanske vlade o upotrebi automatskog sistema kažnjavanja

[https://www.theregister.co.uk/2017/02/08/ministry\\_justice\\_conviction\\_by\\_computer\\_scheme\\_go\\_ahead/](https://www.theregister.co.uk/2017/02/08/ministry_justice_conviction_by_computer_scheme_go_ahead/)

10. Amazon odbacuje svoj alat veštačke inteligencije koji je koristio za zapošljavanje

<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

11. ProPublica je objavio analizu o pristrasnosti COMPAS

<https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

12. Autonomna vozila

[https://en.wikipedia.org/wiki/Self-driving\\_car](https://en.wikipedia.org/wiki/Self-driving_car)