



Piše: Branko Pavlović, predsednik Udruženja aktuara Srbije

■ U delatnosti osiguranja u svetu su već realizovani brojni projekti bazirani na mašinskom učenju u borbi protiv prevara, kao i određivanju tarifa i rešavanju šteta, naročito u zdravstvenom osiguranju i osiguranju vozila. Koliko mi je poznato, već se radi na sličnim projektima i u domaćim osiguravajućim kompanijama i verovatno će se o tome više čuti u narednim mesecima

Udruženje aktuara Srbije i Ekonomski fakultet Univerziteta u Beogradu organizovali su u maju na Zlatiboru XVII međunarodni simpozijum sa temom "Osiguranje na pragu IV industrijske revolucije". Bilo je puno zanimljivih predavanja o raznim aspektima savremenih informatičkih trendova u osiguranju, o čemu će biti više reči u drugom tekstu u ovom broju časopisa. Na Simpozijumu sam prezentovao rad o primeni softverskih robota u osiguranju. Pošto sam o toj temi već pisao u ovoj kolumni pre nekoliko meseci, sada ću se posvetiti jednom podjednako inspirativnom aspektu IV industrijske revolucije, sa velikim potencijalom u primeni u osiguranju - mašinskom učenju. To je podskup veštačke inteligencije namenjen za proučavanje i prepoznavanje obrazaca u ponašanju, korišćenjem statističkih metoda obrade dostupnih podataka. Drugim rečima mašinsko učenje je softver koji može

# Mašinsko učenje za početnike

samostalno da zaključuje na osnovu prethodnog iskustva. Ključna korist koju kompanije mogu da imaju od mašinskog učenja je predviđanje budućih trendova, na osnovu toga što softver samostalno uoči obrasce u dostupnim strukturiranim, ali i nestrukturiranim podacima.

## Koncept

Mašinsko učenje se zasniva na mogućnosti specijalizovanog softvera da napravi sopstvenu logiku i samostalno uoči veze između podataka. Jedan od realizovanih primera mašinskog učenja u praksi je klasifikacija poruka elektronske pošte na spam ili regularnu poštu.

Veliki broj problema iz prakse se može svesti na neku vrstu funkcije koja zavisi od određenog broja parametara. Npr. cene nekretnine zavise od površine, broja soba, lokacije, godine gradnje, vrste grejanja, sprata, itd. Ako softver nekako uspe da poveže sve te podatke ili samo podatke koje on izabere, tako da može da odredi cenu nove nekretnine, smatra se da je softver "naučio" da radi procenu cene. Kod mašinskog učenja, najbolje je što čovek ne mora da zna ništa o tome kako je softver odredio funkciju međuzavisnosti podataka. Uloga čoveka je da obezbedi dovoljno podataka iz prošlosti npr. o prodatim nekretninama, a uloga softvera je da te podatke nekako poveže.

Ipak, bitno je razumeti kako to softver radi. Prvo pokuša da uspostavi uprošćenu funkciju cene nekretnine pomoću npr. 3 promenljive, na sledeći način:

$$\text{Cena} = \text{površina} * x_1 + \text{lokacija} * x_2 + \text{starost} * x_3 + x_4$$

$x_1$ - $x_4$  su tzv. težinski faktori koje softver bira tako da se dobije optimalno rešenje, tj. da se dobije najmanja greška. Na početku može da svakom težinskom faktoru dodeli vrednost 1 i pusti sve primere prodatih nekretnina koje ima. Naravno, dobijene cene na osnovu ove funkcije će se mnogo razlikovati od originalnih. Kada se sabere sve greške na kvadrat, dobija se ukupna greška.

Sledeći korak je menjanje težinskih faktora s ciljem da ukupna greška bude minimalna. Ukoliko bi se postiglo da je ukupna greška jednaka nuli, to bi značilo da je funkcija savršena i da se svi primeri iz

prošlosti potpuno mogu opisati datom funkcijom i izabranim težinskim faktorima, odnosno da se sa velikom verovatnoćom, na osnovu pomenutih parametara (površina, lokacija i starost) može pogoditi i cena za koju će se prodati sledeća nekretnina. Naravno, menjanje težinskih faktora se ne vrši nasumičnim pogađanjem, nego tako što se minimizira funkcija ukupne greške pomoću parcijalnih izvoda.

U praksi standardni softver za mašinsko učenje sve to sam uradi, isprobavajući mnogo komplikovanije međuzavisnosti promenljivih nego što je linearna funkcija data u primeru, ali važno je primetiti da je princip isti.

## Osobine

Iako je koncept mašinskog učenja prilično jednostavan, potrebno je prilično iskustvo da bi se uspešno primenilo na određeni problem iz prakse.

Istraživanja su pokazala da softver za mašinsko učenje koga konfiguriraju iskusni informatičari daje mnogo bolje međuzavisnosti podatka nego kada stručnjak iz određene oblasti pokušava da formuliše pravila na osnovu znanja i iskustva.

Softver daje funkciju, koja je za čoveka koji ga je konfigurisao, crna kutija, ali iako čovek ne razume međuzavisnosti između podataka, može da pokaže na primerima da su rezultati koje daje softver ispravni.

Mašinsko učenje funkcioniše samo u slučajevima kada u realnosti postoji veza između podataka. Ukoliko se u softver ubace podaci o tome koje knjige je pročitao vlasnik nekretnine i cena za koju je nekretnina prodana, nema te magije koja može na osnovu knjiga da odredi cenu nekretnine. Najbolje je uključiti mašinsko učenje u probleme koje bi i čovek mogao rešiti, ali će ih softver za mašinsko učenje rešiti znatno brže i kvalitetnije.

## Primena u osiguranju

U delatnosti osiguranja u svetu su već realizovani brojni projekti bazirani na mašinskom učenju u borbi protiv prevara, kao i određivanju tarifa i rešavanju šteta, naročito u zdravstvenom osiguranju i osiguranju vozila. Koliko mi je poznato, već se radi na sličnim projektima i u domaćim osiguravajućim kompanijama i verovatno će se o tome više čuti u narednim mesecima. ■

