

UDK: 368:371.694:61:659.24:681.3.064:004.4:004.021:679.058:681.322.2:681.3.06
(4-672EEZ) (73)(479.11)
doi: 10.5937/tokosig1903007P

Branko R. Pavlović¹

IZAZOVI U PRIMENI MAŠINSKOG UČENJA U DELATNOSTI OSIGURANJA

PREGLEDNI RAD

Apstrakt

Internet, inteligentni uređaji, Blockchain tehnologija,² softverski roboti³ i razni aspekti veštačke inteligencije, kao što je mašinsko učenje, vrlo često se pominju u savremenoj literaturi kao veliki potencijal za unapređenje procesa u oblasti osiguranja. Ovaj rad biće posvećen mašinskom učenju.

Mašinsko učenje 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 samostalno da zaključuje na osnovu prethodnog iskustva. Ključna korist koju kompanije mogu imati od mašinskog učenja jeste predviđanje budućih trendova, na osnovu toga što softver samostalno uočava obrasce u dostupnim podacima.

Mašinsko učenje može se koristiti u osiguravajućim kućama za optimalnije određivanje tarifa, efikasnije rešavanje i adekvatnije rezervisanje šteta, te moćno oružje u borbi protiv prevara.

Ključne reči: mašinsko učenje, osiguranje

I. Uvod

Mašinsko učenje je metoda analize podataka, koja pomoću specifičnih algoritama omogućava pronalaženje skrivenih veza u podacima, bez eksplicitnog programiranja postupka pronalaska tih veza. Drugim rečima, mašinsko učenje je

¹ Predsednik Udruženja aktuara Srbije
I-mejl: pavlovic.branko@gmail.com
Rad je primljen: 25. juna 2019.
Rad je prihvaćen: 16. jula 2019.

² Pavlović, B. (2018). Blockchain Technology in Insurance and Reinsurance, XVI međunarodni simpozijum „Novi izazovi na tržištu osiguranja“, Arandelovac

³ Pavlović, B. (2019). Robot Usage in Insurance, XVII međunarodni simpozijum „Osiguranje na pragu IV industrijske revolucije“, Zlatibor

sposobnost kompjutera da pronađe nove informacije u odnosu na one što su bile poznate u trenutku učenja.

Primenjuje se u prepoznavanju lica, prepoznavanju objekata na slikama ili na video-klipu (npr. prepoznavanju anomalija na rendgenskim snimcima), zatim u autonomnoj vožnji automobila, igranju igara na tabli (npr. šah, dame, tavlja i go), igranju računarskih igara (npr. Doom i Counter Strike), u kvizovima, klasifikaciji teksta, prevodenju, analizi društvenih mreža, prepoznavanju govora itd.

Predmet rada jeste analiza mogućnosti za unapređenje procesa u delatnosti osiguranja korišćenjem tehnika mašinskog učenja, kao i izazova koji se pritom javljaju. Cilj rada jeste da ukaže na koristi od uključivanja srpskih osiguravajućih kuća u jedan od trendova koji donosi IV industrijska revolucija.

1. Podela metoda mašinskog učenja

Mašinsko učenje zasniva se na mogućnosti specijalizovanog softvera da napravi sopstvenu logiku i samostalno uoči veze između podataka. Koriste se tri metoda⁴ rešavanja problema.

1.1. Nadgledano učenje

Podrazumeva formiranje trening skupa podataka i njegovo korišćenje za učenje softvera. Uočene veze između podataka u trening skupu softver primenjuje na novi skup podataka, koji do tada nije video. Primer je klasifikacija tkiva na maligna i benigna.

Problemi koji se rešavanju metodom nadgledanog učenja dele se na regresivne i klasifikacione.

Kod regresivnih problema cilj je da se unete promenljive povežu i da se rezultat predstavi neprekidnom funkcijom. Primer je predviđanje cena nekretnina na osnovu prethodnog iskustva praćenja transakcija na tržištu nekretnina.

Kod klasifikacionih problema, cilj je da se rezultat predstavi u obliku diskretne izlazne vrednosti. Primer je prepoznavanje rukopisa.

Sličnosti i razlike regresivnih i klasifikacionih problema mogu se objasniti na sledećem primeru:⁵ date su realne veličine kuća zajedno s njihovim cenama u trening skupu i cilj je da se na osnovu veličine kuće predvidi njena cena. Cena kao funkcija u zavisnosti od veličine kuće predstavlja neprekidni izlazni parametar. Dakle, to je regresivni problem. Ali ukoliko je cilj da se sazna da li je kuća prodana za manju ili veću cenu od predviđene, umesto da se predvidi cena kuće, to postaje klasifikacioni problem. Tada se kuće klasifikuju u dve diskretne kategorije u zavisnosti od toga da li su prodane po predviđenoj ceni (jedna kategorija) ili nisu (druga kategorija).

⁴ Nikolić, M. i Zečević, A. (2019). Mašinsko učenje, Matematički fakultet Univerziteta u Beogradu

⁵ Milovanović, A. (2015). Seminarski rad „Mašinsko učenje“, Matematički fakultet Univerziteta u Beogradu

1.2. Nenadgledano učenje

Ne koristi se trening skup, nego samo tehnike pronalaska skrivenog znanja na skupu koji se analizira. Primer⁶ je klasifikacija tkiva u grupe sa sličnim karakteristikama.

Kod nenadgledanog učenja mogu da se formiraju strukture podataka (grupe ili klasteri). Klasteri se prave na osnovu povezanosti promenljivih u podacima. Cilj je da se uoče pravilnosti među podacima i nema povratne informacije zasnovane na predviđenim rezultatima, to jest ne zna se šta je tačno, a šta nije. Zbog toga se potencijalno rešenje problema ovim metodom ne može optimizovati. Klasterovanje predstavlja dobar metod za, na primer, tagovanje dokumenata.

1.3. Učenje potkrepljivanjem⁷

Teorijski okvir učenja potkrepljivanjem opisuje se Markovljevim procesima odlučivanja (engl. Markov Decision Processes). Neka agent i okruženje imaju interakciju samo u diskretnim trenucima $t = 0, 1, \dots$. U svakom trenutku t , agent opaža stanje okruženja S_t iz konačnog skupa stanja S i preduzima akciju A_t iz konačnog skupa dopustivih akcija $A(s)$ u konkretnom stanju „ s “, dobija nagradu R_{t+1} iz konačnog skupa nagrada R i prelazi u novo stanje S_{t+1} . Osnovno je Markovljevo svojstvo da novo stanje i nagrada zavise samo od prethodnog stanja i preduzete akcije, a ne od cele istorije procesa.

Nagrade implicitno definišu cilj agenta. Određuju se tako da pokušavajući da maksimira nagradu, agent obavlja posao koji treba da se obavi. Klasična greška je da se nagrade koriste za usmeravanje agenta kako da nešto uradi, umesto šta da uradi. Na primer, ukoliko agent igra šah, nagradu treba da dobije ukoliko pobedi u partiji, a ne ukoliko pojede protivničku figuru.

II. Pregled literature

Veliki broj autora u svetu bavio se mašinskim učenjem u poslednjih sedamdeset godina, dok se su primenom ove metode analize podataka u delatnosti osiguranja autori počeli baviti relativno skoro. Artur Samjuel dao je neformalnu definiciju mašinskog učenja⁸ 1959. godine, na sledeći način: „Oblast nauke koja daje računaru mogućnost da uči a da prethodno ne bude eksplicitno programiran.“ Savremenu naučnu definiciju⁹ dao je Tom Mičel 1997. godine: „Kaže se da računarski program uči iz iskustva E , uzimajući u obzir određenu klasu zadataka T i merilo učinka P , ako se njegov učinak u zadacima T , meren sa P , poboljšava sa iskustvom E .“

U Tabeli 1 su prikazani najznačajniji događaji u razvoju mašinskog učenja u prethodnom periodu.

⁶ Novaković, J. (2013). Rešavanje klasifikacionih problema mašinskog učenja, Reinženjering poslovnih procesa, vol. 4, Fakultet tehničkih nauka u Čačku

⁷ Nikolić, M. i Zečević, A. (2019). *Mašinsko učenje*, Matematički fakultet Univerziteta u Beogradu

⁸ Samuel, L. (1959). Some studies in machine learning using the game of checkers, *IBM Journal of research and development*, 3(3), p. 210–229.

⁹ Mitchell, T. (1997). *Machine Learning*, McGraw Hill and MIT Press, USA

Tabela 1. Pregled istorijskog razvoja mašinskog učenja

| Godina | Događaj |
|---------------|---|
| 1943. | Mekaloh i Pits su formulisali „threshold logic“, preteču neuronskih mreža |
| 1950. | Alen Tjuring se bavio mašinama koje uče |
| 1952. | Artur Semjuel je napisao prvi program koji uči da igra igru dame |
| 1957. | Frenk Rozenblat je napravio hardverski „perceptron“ |
| 1963. | Vapnik i Červonenkis su predložili prvu varijantu metoda potpornih vektora |
| 1967. | Kaver i Hart su predložili algoritam „k najbližih suseda“ s primenom u problemu trgovačkog putnika |
| 1975. | Verbos je formulisao algoritam propagacije unazad (engl. back-propagation) za izračunavanje gradijenta neuronske mreže |
| 1981. | Dedžong je uveo učenje zasnovano na objašnjavanju kojim se omogućuje izvođenje pravila iz podataka |
| 1985. | Sejnovski i Rozenberg napravili su softver koji može da uči da izgovara engleske reči |
| 1992. | Bozer, Gijon i Vapnik predložili su upotrebu kernela s metodom potpornih vektora, što je dovelo do toga da ovaj metod dominira oblašću mašinskog učenja tokom devedesetih |
| 1992. | Tezauro je napravio TD-Gammon, sistem koji igra igru tavlja (engl. Backgammon) |
| 2006. | Hinton je uveo izraz duboko učenje (engl. deep learning) za algoritme za trening višeslojnih neuronskih mreža koje od tada dominiraju oblašću mašinskog učenja |
| 2011. | IBM-ov sistem Watson pobedio je prvake u američkom popularnom televizijskom kvizu Jeopardy! |
| 2012. | Google X je razvio sistem koji je u stanju da sam pregleda video zapise na YouTube-u i prepoznaje mačke |

| Godina | Događaj |
|---------------------|--|
| 2014. ¹⁰ | Facebook sistem za prepoznavanje lica pomoću neuronskih mreža, DeepFace, dostigao je uspešnost prepoznavanja oko 97% |
| 2016. | Googleov sistem Alfa Go pobedio je svetskog prvaka u igri go rezultatom 4:1 |
| 2017. ¹¹ | Alphabetov sistem Jigsaw, zasnovan na mašinskom učenju, uspeo je da prepozna i zaustavi poplavu nerelevantnih, automatskih komentara koje su ostavljali roboti na sajtovima (engl. on-line trolling) |
| 2018. ¹² | Unapređen model prepoznavanja teksta govornih jezika pomoću neuronskih mreža uvođenjem prethodnog treniranja (engl. pretrained language model) |

Izvor: Nikolić, M. (2018). *Uvod u nadgledano mašinsko učenje*, Matematički fakultet Univerziteta u Beogradu

U delatnosti osiguranja autori su se pretežno bavili primenom mašinskog učenja u određivanju tarifa, rezervisanju šteta i borbi protiv prevara. Grupa američkih autora¹⁰ 2019. godine bavila se sledećim tehnikama mašinskog učenja: *simple decision tree*, *random forest* i *boosted trees* na primeru formiranja tarifa u osiguranju na osnovu podataka o štetama. Švajcarski profesor Wuthrich istražio je 2016. godine rezervisanje šteta pomoću Chain Ladder metode analiziranjem ponašanja pojedinačnih šteta metodom mašinskog učenja *regression trees*.¹¹ Grupa australijskih autora¹² je 2018. godine, korišćenjem LASSO modela, opisala potpuno automatizovano rezervisanje šteta na primeru nematerijalnih šteta osiguranja od auto-odgovornosti. Grupa turskih autora¹³ je 2015. godine dala doprinos u formiranju sistema za prevenciju prevara u zdravstvenom osiguranju pomoću metoda mašinskog učenja.

Domaći autori se nisu do sada bavili primenom mašinskog učenja u delatnosti osiguranja u Srbiji.

¹³ Henckaerts, R., Cote, M.-P., Antonio, K. and Verbelen, R. (2019). *Boosting Insights in Insurance Tariff Plans with Tree-based Machine Learning*, Cornell University arXiv:1904.10890, New York, USA

¹¹ Wuthrich, M. (2016). *Machine Learning in Individual Claims reserving*, *Swiss Finance Institute Research Paper Series No. 16–67*, Geneva, Swiss

¹² McGuire, G., Taylor, G. and Miller, H. (2018). *Self-assembling Insurance Claim Models Using Regularized Regression and Machine Learning*, <http://dx.doi.org/10.2139/ssrn.3241906>, Sydney, Australia

¹³ Kose, I., Gokturk, M. and Kilic, K. (2015). *An Interactive Machine-learning-based Electronic Fraud and Abuse Detection System in Healthcare Insurance*, *Applied Soft Computing* 36, p. 283–299

III. Koncept

1. Princip rešavanja problema mašinskim učenjem

Veliki broj problema iz prakse može se 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 odrediti cenu nove nekretnine, smatra se da je softver naučio da procenjuje cenu. Kod mašinskog učenja, 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 princip po kome to softver radi. Jedan od jednostavnijih metoda jeste da prvo pokuša uspostaviti uprošćenu funkciju cene nekretnine pomoću npr. 3 promenljive, na sledeći način:

$$\text{cena}(x_1, x_2, x_3, x_4) = \text{površina} * x_1 + \text{lokacija} * x_2 + \text{starost} * x_3 + x_4$$

gde su x_1 - x_4 tzv. težinski faktori koje softver bira tako da se dobije optimalno rešenje, tj. da se dobije najmanja greška. Na početku svakom težinskom faktoru može da dodeli vrednost 1 i pusti sve primere prodatih nekretnina koje ima. Naravno, dobijene cene na osnovu ove funkcije mnogo će se razlikovati od originalnih. Kada se saberu greške na kvadrat za sve postojeće podatke o prodatim nekretninama, dobija se ukupna greška, koja se može zapisati na sledeći način:

$$E = \sum (\text{stvarna_cena} - \text{cena}(x_1, x_2, x_3, x_4))^2$$

Sledeći korak jeste menjanje težinskih faktora s ciljem da ukupna greška bude minimalna. Ukoliko bi se postiglo to da ukupna greška bude jednaka nuli, to bi značilo da je funkcija savršena i da se svi primeri iz prošlosti mogu potpuno opisati datom funkcijom i izabranim težinskim faktorima, odnosno da se s 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 ne vrši se nasumičnim pogađanjem, nego tako što se minimizira funkcija ukupne greške pomoću parcijalnih izvoda po svakom težinskom faktoru:

$$\frac{\partial E}{\partial x} = 0$$

za x_1, \dots, x_4 .

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 je princip sličan.

2. Osobine

Iako je koncept mašinskog učenja prilično jednostavan, potrebno je poprilično iskustvo da bi se uspešno primenio na određeni problem iz prakse. Istraživanja su pokazala da softver za mašinsko učenje koji konfiguriše iskusni informatičar daje mnogo bolje međuzavisnosti podatka nego kada stručnjak iz određene oblasti pokušava formulisati pravila na osnovu znanja i iskustva.

Softver daje funkciju, koja je za čoveka što 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 je knjige pročitao vlasnik nekretnine i cena za koju je nekretnina prodana, nema te magije koja na osnovu knjiga može da odredi cenu nekretnine. Najbolji rezultati dobijaju se kada se uključi mašinsko učenje u probleme koje bi i čovek mogao rešiti, jer ih softver za mašinsko učenje rešava znatno brže i kvalitetnije.

3. Primena mašinskog učenja

Mašinsko učenje u praksi se svodi na interaktivni proces sa sledećim koracima:

- Obezbeđenje relevantnih podataka za rešavanje datog problema
- Priprema podataka za analizu mašinskim učenjem, s ciljem da se radi samo kvalitetnim i pouzdanim podacima
- Izbor odgovarajućeg algoritma mašinskog učenja
- Trening algoritma, koji može biti nadgledan, nenadgledan i potkrepljen neophodan je za formiranje kvalitetnog modela
- Evaluacija modela, da bi se izabrao algoritam s najboljim performansama na konkretnom problemu
- Kreirani model distribuira se korisnicima u vidu aplikacije
- Korisnici rešavaju problem na osnovu sopstvenih podataka, koje model do tada nije video
- Procena validnosti rešenja problema i zatim vraćanje na početak procesa dok se ne postigne dovoljno dobro rešenje.

4. Algoritmi¹⁴

Algoritmi mašinskog učenja drugačiji su od standardnih algoritama koji se sreću u informatici, jer umesto da algoritam počinje unosom podatka, kod

¹⁴ Hurwitz, J. and Kirsch, D. (2018). *Machine Learning for Dummies*, IBM Limited Edition, John Wiley & Son, New York, USA

mašinskog učenja podaci kreiraju model. Glavna prednost tog pristupa jeste ta što eliminiše predrasude i pogrešne pretpostavke iz modela. Izbor odgovarajućeg algoritma nije lak i potrebno je veliko iskustvo kako bi se napravio dovoljno dobar izbor, koji će dati prilagođen model konkretnom problemu. Najčešće su algoritmi napisani na jednom od sledećih programskih jezika: Java, Python i R.

Najinteresantniji algoritmi mašinskog učenja su sledeći:

- Linearna regresija – to je algoritam koji se najviše koristi u mašinskom učenju. Koristi se za statističke analize tako što uspevaju da uspostave veze između podataka. Algoritmi koji koriste regresiju mogu da kvantifikuju jačinu korelacije između promenljivih u datom skupu podataka, kao i da predvide vrednosti promenljivih u budućnosti, na osnovu podatka o njima iz prošlosti.
- Klasterovanje – objekti sa sličnim osobinama grupišu se u klustere tako da svi objekti u klasteru budu međusobno sličniji u odnosu na objekte iz drugih klastera. Algoritam prvo prepoznaje parametre objekta, a zatim na osnovu parametara razvrstava objekte.
- Stablo odlučivanja – koristi se struktura stabla da se prikažu rezultati odlučivanja. Svaki krajnji čvor stabla predstavlja mogući ishod algoritma, dok iz čvorova odluke izlaze grane u zavisnosti od vrednosti parametra koji se nalazi u čvoru. Osim navedenih, koriste se i algoritmi bazirani na instancama, regularizacija, mašinsko učenje bazirano na pravilima, neuronske mreže, itd.

5. Preduslovi za intenzivan razvoj primene mašinskog učenja

Iako su osnovni principi mašinskog učenja razvijeni pre više od pedeset godina, primena u praksi naročito je intenzivirana u poslednjih nekoliko godina, jer su tek tada ispunjeni potrebni preduslovi.

Moderni računarski procesori dostigli su potrebnu snagu da rade dovoljno brzo s velikom količinom podataka. Troškovi skladištenja podataka dramatično su pali. Pojavile su se relativno jednostavne mogućnosti za povezivanje velikog broja kompjutera u klustere koji mogu analizirati velike količine podataka zadovoljavajućom brzinom. Zahvaljujući opštoj digitalizaciji, mnogo različitih podataka (meteorološki, medicinski, društvene mreže itd.) postaju komercijalno upotrebljivi putem usluga u oblaku i API interfejsima. Algoritmi mašinskog učenja postali su besplatni i lako dostupni zainteresovanima preko open-source zajednice (npr. Googleov TensorFlow).

Zahvaljujući ispunjenim navedenim preduslovima, intenzivna primena mašinskog učenja nije zaobišla ni delatnost osiguranja.

IV. Primena mašinskog učenja u osiguranju

U delatnosti osiguranja u svetu su već realizovani brojni projekti bazirani na mašinskom učenju.

Oko trećina evropskih osiguravajućih kuća koristi mašinsko učenje u

svom poslovanju, pokazala je analiza¹⁵ Evropskog nadzornog tela za osiguranje i profesionalna penzijska osiguranja (engl. *European Insurance and Occupational Pensions Authority* = EIOPA) o upotrebi Big Data Analytics na 28 evropskih tržišta osiguranja u osiguranju od auto-odgovornosti i zdravstvenom osiguranju, u kome su učestvovala 222 entiteta iz delatnosti osiguranja. Metode i softver mašinskog učenja već koristi 31 odsto osiguravajućih kuća, dok još 24 odsto radi na njihovom razvoju. Pomenute napredne metode omogućavaju precizno procenjivanje raznih trendova, sa ljudskom intervencijom ili bez nje, i na taj način povećavaju efikasnost u donošenju poslovnih odluka i smanjuju troškove.

Analiza koju je sproveda EIOPA pokazala je da se tradicionalni izvori podataka (demografski podaci, podaci o izloženosti rizicima i podaci o štetama) češće koriste u kombinaciji s tehnologijom mašinskog učenja nego novi izvori podatka (npr. podaci iz telematike, genetski podaci, podaci o korišćenju kreditnih kartica). Najvažniji rezultat korišćenja tradicionalnih podataka na nov način jesu usluge bolje prilagođene pojedinačnim osiguranicima i valjanija procena rizika. Takođe, osiguravači sve više koriste podatke koje dobijaju od trećih strana za izračunavanje kreditnih bodova, rejtinga na osnovu sigurnosti vožnje, bodovanja višestrukih štetnika itd.

1. Oblasti primene mašinskog učenja u osiguranju

Rezultat istraživanja EIOPA jeste da se mašinsko učenje može koristiti u sledećim oblastima¹⁶ u osiguranju:

- Razvoj novih usluga – uvođenje usluga koje se baziraju na upotrebi određenog resursa (npr. vožnja vozila), usluge skrojene po potrebama pojedinačnog osiguranika, nove vrste osiguranja (npr. cyber osiguranje) itd.
- Tarifiranje i preuzimanje rizika – poboljšana procena rizika, uvođenje novih dodatnih faktora za obračun cene polise, prediktivni modeli za štete, optimizacija cena, modeliranje raskida polisa itd.
- Prodaja – automatizovano savetovanje, izbacivanje posrednika iz procesa prodaje polisa, sofisticirani sistem upravljanja odnosom s klijentima (engl. *Customer Relationship Management*), povećanje broja interakcija s osiguranicima itd.
- Postprodaja – pametne aplikacije na mobilnim uređajima, usluga klijentima tokom svih 24 sata dnevno, *Chat botovi*, automatska obaveštenja o opasnostima od npr. poplave, oluje itd.
- Rešavanje šteta – napredna analitika u prevenciji prevara u osiguranju, prepoznavanje oblika na fotografijama za podršku proceni šteta na vozilima, automatsko prepoznavanje faktura npr. iz servisa za popravku vozila i njihovo plaćanje itd.

¹⁵ EIOPA (2019). Big Data Analytics in Motor in Health Insurance: A Thematic Review, www.eiopa.europa.eu

¹⁶ EIOPA (2019). Big Data Analytics in Motor in Health Insurance: A Thematic Review, www.eiopa.europa.eu

2. Izazovi u primeni mašinskog učenja u osiguranju

Uvođenje mašinskog učenja u delatnost osiguranja može da donese više štete nego koristi ukoliko se ne obrati pažnja na sledeće izazove:

- Diskriminacija – budući da se učenje obično izvodi nad podacima iz prošlosti, kao i da se ne može zabraniti algoritmima mašinskog učenja da pronađu očigledne veze koje su zabranjene za korišćenje prilikom određivanja tarifa po regulativi Evropske unije (npr. da žene žive duže), mora se na kraju proveriti da li je ono što preporučuje softver u skladu s propisima.
- Netransparentnost modela – kao što je već objašnjeno, rad softvera koji primenjuje mašinsko učenje predstavlja crnu kutiju za korisnike, što može izazvati probleme u režimu Solventnost II, koji zahteva da modeli budu transparentni i da ih mogu proveravati revizori.
- Nefer poslovanje – optimizacija cena može rezultovati time da softver u obračun uključi ne samo faktore rizika nego i faktore o osetljivosti pojedinih grupa na promenu cene ili nesklonost drugih grupa da pretražuju tržište u cilju pronalaska najbolje ponude. Takođe, softver bi mogao da pronađe koliki je procenat iznosa pune štete većina oštećenih spremna da prihvati i da odustane od sudskog postupka za potraživanje ostatka iznosa. Ovakva praksa ne bi bila etična i takav poslovni odnos s osiguranicima ne bi se mogao uzeti kao fer odnos.
- Zloupotreba genetskih podataka – s obzirom na to da genetski podaci osiguranika mogu postati dostupni preko eksternih kompanija, softver ih može upotrebiti za određivanje cena zdravstvenog ili životnog osiguranja, što nije u skladu s osiguravajućom praksom u Evropskoj uniji.
- Neusaglašenost sa GDPR-om – kompanije koje koriste crnu kutiju mašinskog učenja teško mogu dokazati da su usaglašene sa Opštom uredbom o zaštiti podataka o ličnosti Evropske unije (engl. *General Data Protection Regulation* = GDPR), odnosno da na pravi način rukuju ličnim podacima, što u Evropskoj uniji može dovesti do velikih novčanih kazni.
- Kompromitovanost podataka – cyber rizici su sve veći u savremenom poslovanju, što može ishodovati situacijom da softver mašinskog učenja koristi podatke koji su kompromitovani, pa izvede neodgovarajuće zaključke, a da ljudi koji upravljaju njime toga uopšte nisu svesni.
- Neadekvatnost klasičnih polisa¹⁷ – ako softver sa sposobnošću mašinskog učenja prouzrokuje štetu, postavlja se pitanje ko je za to odgovoran. Pitanje osiguranja od odgovornosti za upotrebu veštačke inteligencije je komplikovano, jer potencijalni krivci mogu biti od programera algoritama, preko koda i integratora, do vlasnika skupa podataka i proizvođača koji sve to koristi u svom proizvodu. Skriveni rizici koje nosi veštačka inteligencija nateraće osiguravače da u budućnosti unaprede svoje polise.

¹⁷ Papović, B. (2019) Prikaz članka „Veštačka inteligencija – skriveni rizici“, izvor *Insurance Post*, novembar 2018, 37–38 str, Tokovi osiguranja 1/2019, p. 107–108

V. Primeri primene mašinskog učenja u osiguranju u praksi

Jedan od realizovanih primera mašinskog učenja u praksi, koji se koristi i u osiguravajućim kućama jeste klasifikacija poruka elektronske pošte na spam ili regularnu poštu.

Ipak, postoji i nekoliko realizovanih primera koji su specifični samo za oblast osiguranja u kojima je ova tehnologija primenjena. Četiri najveće osiguravajuće kuće u SAD, po zvaničnim podacima¹⁸ iz 2016. godine, koriste mašinsko učenje u različitim poslovnim procesima¹⁹, dok u Srbiji Udruženje osiguravača Srbije koristi algoritme mašinskog učenja u borbi protiv prevara.

1. State Farm²⁰

Najveća američka osiguravajuća kuća u 2016. godini „State Farm“ koristi mašinsko učenje u klasifikaciji sigurnosti vožnje svakog vozača koji je osiguran kod njih, a zatim i u pripremi ponude odgovarajućih usluga za svaku klasu vozača.

Pre tri godine „State Farm“ otvorila je konkurs za najuspešnije rešenje za klasifikaciju vozača, na osnovu fotografija iz saobraćaja. Vozači su klasifikovani u deset kategorija: sigurni vozači, oni koji šalju SMS poruke u vožnji, vozači koji podešavaju radio tokom vožnje, oni koji razgovaraju preko telefona itd. Prvo mesto je osvojilo rešenje koje je koristilo mašinsko učenje i metodu dveju neuronskih mreža.

2. Liberty Mutual²¹

Jedna od najvećih američkih osiguravajući kuća, „Liberty Mutual“, osnovala je „Solaria Labs“, specijalizovanu kompaniju za inovacije. Jedan od projekata koji je „Solaria Labs“ razvila 2017. godine jeste razvoj portala s otvorenim aplikacionim programskim interfejsom ili API interfejsom (skraćeno od engleskog naziva *Application Programming Interface*). Svrha portala jeste integrisanje sopstvenih informatičkih projekata „Solaria Labsa“ s javnim podacima, s ciljem razvoja aplikacije bazirane na mašinskom učenju za poboljšanje sigurnosti u saobraćaju.

Takođe, razvili su aplikaciju za mobilne uređaje koja pomaže vozačima što su imali saobraćajnu nesreću da brzo procene štetu na vozilu korišćenjem kamere mobilnog telefona. Trening skup aplikacije bazirane na mašinskom učenju sastojao se od više hiljada fotografija saobraćajnih nesreća uparenih s podacima o troškovima popravki vozila sa fotografija.

¹⁸ National Association of Insurance Commissioners web site www.naic.org

¹⁹ Sennaar, K. (2019). How America's Top 4 Insurance Companies are Using Machine Learning, Emerj Artificial Intelligence Research, <http://emerj.com>, Boston, USA

²⁰ www.statefarm.com

²¹ www.libertymutual.com

3. Allstate²²

Osiguravajuća kuća „Allstate“ razvila je virtuelnog asistenta, *Chat-bot* ABIE (skraćeno od engl. *Allstate Business Insurance Expert*), zasnovanog na metodama mašinskog učenja, za pomoć svojim agentima u prodaji proizvoda za korporativne klijente. Ranije su njihovi agenti bili obučeni samo za prodaju osiguranja kuća i stanova i zdravstvenog osiguranja, tako da im je bila potrebna pomoć kada je kompanija odlučila da im poveri prodaju složenijih usluga imovinskih osiguranja. ABIE je obezbedio agentima vodič „korak po korak“ na engleskom jeziku za primenu tarifa, određivanje cena i pripremu ponude raznih usluga osiguranja za korporativne klijente. Budući da je ABIE pružao kvalitetnu pomoć, agenti su koristili njegove usluge u oko 25.000 slučajeva mesečno.

4. Progressive²³

Velika američka osiguravajuća kuća „Progressive“ koristi algoritme prediktivne analitike mašinskog učenja za proučavanje podataka o vozačima s ciljem razumevanja tržišnih trendova i mogućnosti za dalji razvoj usluga osiguranja vozila. Njihova aplikacija za telematiku, Snapshot, u 2016. godini prikupila je podatke o vožnji na preko dvadeset milijardi kilometara.

5. Udruženje osiguravača Srbije²⁴

Skoro sve veće srpske osiguravajuće kuće imaju formirane organizacione delove za sprečavanje prevara, koje se sa većim ili manjim uspehom bore protiv prevara. U cilju što efikasnije borbe s prevarantima iz oblasti osiguranja vozila, Udruženje osiguravača Srbije obezbedilo je 2015. godine osiguravajućim kompanijama moćan alat za borbu protiv prevara. FROPS (engl. *Fraud Risk Operational Performance Solution*) softver britanske kompanije „Salviol“,²⁵ za prevenciju prevara, analitički je alat koji se koristi u osiguranju i drugim finansijskim sektorima. Glavni cilj FROPS-a jeste da sakuplja, upoređuje, istražuje i analizira velike količine podataka. On traži anomalije odstupanja ili nekoherentnosti u podacima s ciljem otkrivanja prevarnih radnji koje dovode do gubitka prihoda. FROPS analizira najširi spektar informacija kako bi obezbedio najefikasnije analitičko okruženje s maksimalnom preciznošću ključnih indikatora prevara. On vrši indeksiranje podataka iz kojih obračunava rizik od prevare. Funkcioniše nezavisno od drugih programskih okruženja. Koristi prediktivnu analitiku i mašinsko učenje za identifikovanje novih tipova prevara u osiguranju.

²² www.allstate.com

²³ www.progressive.com

²⁴ www.uos.rs

²⁵ www.salviol.com

Ključni indikatori prevara jesu parametri štete koji s velikom verovatnoćom ukazuju na postojanje prevarne radnje. Oni predstavljaju kritičnu tačku za uspostavljanje sistema za otkrivanje i sprečavanje prevara. Pravilnim kombinovanjem tih indikatora, sužava se skup šteta koje su potencijalno prevare. Sektor za borbu protiv prevara fokusira se na taj skup u daljoj kontroli i analizi, što povećava njegovu efikasnost. Prepoznavanjem nestrukturiranih podataka, tekstualnih datoteka, skeniranih dokumenta itd. FROPS omogućava njihovo pretraživanje, analizu i kategorizaciju. Tako pripremljeni podaci daju analitičaru neophodnu širinu za proces istrage.

Takođe, FROPS analizira socijalne mreže kvantitativnom tehnikom koja kombinuje organizacionu teoriju s matematičkim modelima. Na osnovu tih algoritma, analitičar može bolje razumeti dinamiku grupa, mreža i organizacija. Može da identifikuje lica ili organizaciju koja krije informaciju, da prati i analizira kako se mreže razvijaju u toku vremena i da velikom brzinom utvrdi važne karakteristike posmatrane grupe.

VI. Zaključak

U delatnosti osiguranja u svetu su već realizovani brojni projekti bazirani na mašinskom učenju u borbi protiv prevara, određivanju tarifa, rešavanju šteta, podršci korisnicima i agentima, naročito u zdravstvenom osiguranju i osiguranju vozila. U Srbiji je najpoznatiji primer upotrebe mašinskog učenja u osiguranju softver FROPS, koji koristi Udruženje osiguravača Srbije.

Doprinos ovog rada je demistifikacija primene mašinskog učenja u osiguranju, uz skretanje pažnje na izazove koje korišćenje mašinskog učenja donosi – sve kao podsticaj domaćim osiguravajućim kuća da unapređenje i osavremenjivanje pojedinih internih poslovnih procesa počnu da vrše korišćenjem tehnika mašinskog učenja.

S obzirom na brojne koristi od uvođenja savremenih tehnologija u delatnost osiguranja, koje su opisane u radu i potkrepljene primerima iz strane i domaće prakse, očekuje se masovnije korišćenje mašinskog učenja u srpskim osiguravajućim kućama u narednih nekoliko godina.

Literatura

- EIOPA (2019). Big Data Analytics in Motor in Health Insurance: A Thematic Review, www.eiopa.europa.eu
- Henckaerts, R., Cote, M.-P., Antonio, K. and Verbelen, R. (2019). *Boosting Insights in Insurance Tariff Plans with Tree-based Machine Learning*, Cornell University arXiv:1904.10890, New York, USA
- Hurwitz, J. and Kirsch, D. (2018). *Machine Learning for Dummies*, IBM Limited Edition, John Wiley & Son, New York, USA
- Kose, I., Gokturk, M. and Kilic, K. (2015). An Interactive Machine-learning-based Electronic Fraud and Abuse Detection System in Healthcare Insurance, *Applied Soft Computing* 36, p. 283–299

- McGuire, G., Taylor, G. and Miller, H. (2018). Self-assembling Insurance Claim Models Using Regularized Regression and Machine Learning, <http://dx.doi.org/10.2139/ssrn.3241906>, Sydney, Australia
- Milovanović, A. (2015). Seminarski rad „Mašinsko učenje“, Matematički fakultet Univerziteta u Beogradu
- Mitchell, T., (1997). *Machine Learning*, McGraw Hill and MIT Press, USA
- National Association of Insurance Commissioners web site www.naic.org
- Nikolić, M. i Zečević, A. (2019). *Mašinsko učenje*, Matematički fakultet Univerziteta u Beogradu
- Novaković, J. (2013). Rešavanje klasifikacionih problema mašinskog učenja, *Reinženjering poslovnih procesa*, vol. 4, Fakultet tehničkih nauka u Čačku
- Papović, B. (2019). Prikaz članka „Veštačka inteligencija – skriveni rizici“, izvor *Insurance Post*, November 2018, 37–38 str, Tokovi osiguranja 1/2019, p. 107–108
- Pavlović, B. (2018). Blockchain Technology in Insurance and Reinsurance, XVI međunarodni simpozijum „Novi izazovi na tržištu osiguranja“, Aranđelovac
- Pavlović, B. (2019). Robot Usage in Insurance, XVII međunarodni simpozijum „Osiguranje na pragu IV industrijske revolucije“, Zlatibor
- Ruder, S. (2018). A Review of the Neural History of Natural Language Processing, www.ruder.io
- Samuel, L. (1959). Some studies in machine learning using the game of checkers, *IBM Journal of research and development*, 3(3), p. 210–229
- Sennaar, K. (2019). How America's Top 4 Insurance Companies are Using Machine Learning, Emerj Artificial Intelligence Research, <http://emerj.com>, Boston, USA
- Taigman, Y., Yang, M., Ranzato, M. and Wolf, L. (2014). *DeepFace: Closing the Gap to Human-Level in Face Verification*, Facebook AI Research, Menlo Park, USA
- Wuthrich, M. (2016). Machine Learning in Individual Claims reserving, *Swiss Finance Institute Research Paper Series No. 16–67*, Geneva, Swiss
- www.allstate.com
- www.jigsaw.google.com
- www.libertymutual.com
- www.progressive.com
- www.salviol.com
- www.statefarm.com
- www.uos.rs