Ekonomske ideje i praksa 57 (2025) 29-45

https://doi.org/10.54318/eip.2025.mr.007

KOMPARATIVNA ANALIZA MODELA KREDITNOG SKORINGA: KONVENCIJALNI VS MODELI BAZIRANI NA MAŠINSKOM I DUBOKOM UČENJU

Milovan Rankov[1], e-mail: milovan.rankov@gmail.com

Apstrakt. *Ovaj istraživački rad predstavlja komparativnu analizu različitih modela kreditnog skoringa, fokusirajući se na poređenje logističke regresije sa naprednim modelima mašinskog i dubokog učenja. Kao indikatori performansi modela i osnova za poređenje njihove efikasnosti biće korišćeni tačnost, preciznost, F1, opoziv, Gini i AUC. Jedan od osnovnih ciljeva ovog rada jeste da na transparentan način prikaže efikasnost različitih model i ukaže na njihove prednosti odnosno nedostatke. Za emprijsku analizu biće korišćena „Kaggle" baza podataka o ponašanju dužnika, a samo modeliranje će biti rađeno u razvojnom okruženju PyCharm koristeći Python programski jezik. Jedan od osnovnih rezultata istraživanja jeste da su modeli bazirani na dubokom učenju daleko efikasniji od ostalih modela, posmatrano kroz prizmu pomenutih indikatora performansi.*

Ključne reči: modeli kreditnog skoringa, logistička regresija, mašinsko učenje, duboko učenje.

COMPARATIVE ANALYSIS OF CREDIT SCORING MODELS: CONVENTIONAL VS MODELS BASED ON MACHINE AND DEEP LEARNING

Abstract. *This research paper presents a comparative analysis of credit scoring models, focusing on comparing logistic regression with models based on machine and deep learning. Accuracy, precision, F1, recall, Gini and AUC will be used as performance indicators and the basis for comparing model efficiency. One of the main goals of this research is to transparently compare different models and point out their advantages and disadvantages. The „Kaggle" database on debtor behaviour will be used for empirical analysis, and the modeling itself will be done in the PyCharm development environment using the Python programming language. One of the main results of the research is that models based on deep learning are far more efficient than other models, viewed through the prism of the mentioned performance indicators.*

Key Words: credit scoring models, logistic regression, machine learning, deep learning.

JEL klasifikacija (classification): C45, G21, G32

[1] Student, Ekonomski fakultet, Univerzitet u Beogradu, Kamenička 6, 11000 Beograd,

1. Uvod

U današnjem izuzetno dinamičnom finansijskom okruženju, koje se sa jedne strane brzo razvija u smislu usluga koje su ponuđene, pružaoca tih usluga i kanala preko kojih su ponuđene, kao i sa aspekta tražnje odnosno potreba pojedinaca i privrede, modeli kreditnog skoringa igraju ključnu ulogu u određivanju kreditne sposobnosti tražioca kredita. Efikasni i precizni modeli kreditnog skoringa su od suštinskog značaja za (i) finansijske institucije kako bi minimizirale rizik i maksimizirale prinose, i za (ii) celokupno društvo, jer doprinose pravičnosti procesa kreditiranja, što dalje znači da se svakom pojedincu (ili pravnom licu) povećava šansa za dostupnost dodatnih finansijskih sredstava ukoliko ima potrebu za njima[2].

Modeli kreditnog skoringa se koriste za procenu kreditne sposobnosti pojedinaca (ili pravnih lica) na osnovu analize različitih relevantnih faktora, kao što su demografske karakteristike (starost, bračni status, i slično), finansijski podaci (plata, ostala dugovanja, posedovanje kreditnih kartica i slično), kreditna istorija i drugi parametri[3]. Kao rezultat ovih modela dobija se kvantitativna procena da li će klijent redovno izmirivati obaveze ili postoji rizik da će doći do kašnjenja ili potpune nelikvidnosti[4].

Motivacija za ovo istraživanje proizilazi iz potrebe da se proceni efikasnost tradicionalnih statističkih modela u odnosu na moderne tehnike mašinskog učenja (ML) i dubokog učenja (DL) u kreditnom skoringu. Identifikovanje modela koji poboljšavaju tačnost predviđanja i poboljšavaju procese donošenja odluka u proceni kreditnog rizika je imperativ, s obzirom na sve veću složenost i obim finansijskih podataka[5].

Razvoj modela kreditnog skoringa je doživeo značajan napredak u poslednjih nekoliko decenija. U početku, banke su se uglavnom oslanjale na jednostavne statisičke metode. Logistička regresija (LR) je bila među prvim statističkim tehnikama primenjenim za procenu skoringa[6].

Uprkos širokoj primeni i prednostima, modeli kreditnog skoringa suočavaju se sa nekoliko izazova. Sa jedne strane tradicionalni statistički modeli (LR) često ne mogu da otkriju kompleksne i nelinearne veze u velikim i raznovrsnim skupovima podataka. Nasuprot tome, ML i DL modeli iako nude bolje performanse, oni su izuzetno složeni i zahtevaju značajne računarske resurse.

[2] Abel, Bernanke i Croushore (2007), p. 59
[3] Božović, M., (2021), p.3-6
[4] Abby, W., (2024) p. 3
[5] Thomas, Edelman i Crook (2004), p. 15-45
[6] Thomas, Edelman i Crook (2004), p. 35-45

Drugi značajan izazov je interpretabilnost ovih modela, što je ključno sa aspekta praktične primene, ali i regulative, odnosno usklađenost sa propisima kontrolnog tela[7].

Ovo istraživanje daje uporednu analizu različitih modela kreditnog skoringa. Konkretno, fokus će biti na komparativnoj analizi statističkog modela - logističke regresije (LR) i kompleksnijih modela mašinskog učenja poput stabla odlučivanja (DT), nasumične šume (RF) i na kraju modela dubokog učenja - duboke neuronske mreže (DNN). Za upoređivanje modela biće korišćeni ključni indikatori učinka kao što su tačnost, preciznost, opoziv, F1 rezultat i površina ispod krive (AUC). Ovo istraživanje nastoji da identifikuje prednosti i slabosti svakog pristupa i pruži uvid u njihovu praktičnu primenu u procesu upravljanja kreditnim rizikom[8].

Ostatak rada je podeljen na sledeća poglavlja. U drugom delu biće predstavljen teorijski i metodološki okvir ove studije. Nakon toga će u poglavlju 3 biti prezentovane metode poređenja performansi. Zatim će u poglavlju 4 biti prezentovane ključne karaterisitke baze podataka, a u poglavlju 5 biće detaljno prezentovani empirijski rezultati, uključujući poređenja performansi na osnovu ključnih pokazatelja.

2. Teorijski i metodološki okvir

2.1 ***Logistička regresija (LR)*** predstavlja jednu od najčešće korišćenih statističkih metoda za rešavanje binarnih problema klasifikacije, odnosno za predviđanje dualnih ishoda. U konkretnom primeru procene kreditnog skoringa, ovi modeli nalaze široku primenu za previđanje ponašanja dužnika, gde postoje dva moguća potencijalna ishoda, redovno vraćanje duga banci ili kritično kašnjenje[9].

Logistička funkcija, preslikava bilo koji broj realne vrednosti u opseg [0, 1], što ga čini pogodnim za procenu verovatnoće. Logistička funkcija je definisana kao:

$$p = \frac{1}{1 + e^{-\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_m}} \#(1)$$

Gde su:

β_0 – konstanta, odnosno osnovna verovatnoća da će se događaj desiti kada su sve nezavisne promenljive nule

[7] Baesens, B., et al. (2003), p. 627-635
[8] Barddal, J. P., et al. (2020), p. 4-6
[9] Agarwal, S., Alok, S., Ghosh, P., & Gupta, S. (2023), p 11-15

β_i – koeficijenti predstavljaju promenu logit zavisne promenljive zavisne promenljive za jednu jedinicu promene odgovarajuće nezavisne promenljive, držeći sve ostale promenljive konstantnim. Pozitivan koeficijent povećava log šanse (a samim tim i verovatnoću), dok ga negativan koeficijent smanjuje

x_j – su prediktori koji se koriste za procenu verovatnoće zavisne promenljive. U kontekstu kreditnog skoringa, ove promenljive mogu uključiti informacije specifične za zajmoprimca kao što su prihod, kreditna istorija, odnos duga i prihoda, status zaposlenja i drugi (ne)finansijski pokazatelji.

Modeli logističke regresije se obično ocenjuju korišćenjem metode maksimalne verovatnoće (MLE), koja pronalazi vrednosti parametara β_0, β_1, ... β_n koji maksimiziraju verovatnoću realizacije datih podataka. Funkcija verovatnoće se konstruiše iz logističke funkcije i posmatranih ishoda[10].

Jedna od ključnih prednosti LR je interpretabilnost. Koeficijenti se mogu direktno tumačiti u smislu odnosa verovatnoće i pojedinačnih promenljivih, pružajući uvid u važnost i uticaj svake od njih. Na primer, koeficijent od 0,5 za dohodak znači da povećanje prihoda za jednu jedinicu povećava logaritam šanse za neizvršenje obaveza za 0,5, pod pretpostavkom da se sve ostale promenljive drže konstantnim.

Međutim ta jednostavnost i lakoća tumačenja, dolaze sa nekim od važnih nedostataka ovog modela, kao što su pretpostavka nelinearnog odnosa između zavisne i nezavnih promenljivih zasnovanog na logističkoj funkciji, što vrlo često nije slučaj. Pored toga, performanse LR u velikoj meri zavise od kvaliteta i inžinjeringa promenljivih.

2.2 Stabla odlučivanja (DT) su modeli koji dele podatke u podskupove na osnovu vrednosti ulaznih karakteristika. Proces donošenja odluka predstavljen je kroz strukturu nalik stablu gde svaki čvor predstavlja osobinu (atribut), svaka grana predstavlja pravilo odlučivanja, a svaki list predstavlja ishod[11].

Proces funkcioniše tako što se odabira karakteristika koja najbolje klasifikuje podatke u različite grupe. Na osnovu odabranih karakterisitka, skupovi podataka se dalje dele na podskupove. Ovaj proces se rekurzivno ponavlja za svaki podskup dok se ne ispuni neki od definisanih uslova kao što su maksimalna dubina stabla ili minimalni broj uzoraka po listu[12].

[10] Dumitrescu, L., et al. (2022), p. 1182-1186
[11] Brieman, et al. (1984), p. 280
[12] Chang, A.-H., et al. (2022), p.308

Gini indeks (ili Gini indeks čistoće) meri (ne)čistoću samog čvora:

$$Gini = 1 - \sum_{i=1}^{C} p_i^2 \#(2)$$

Gde je p_i verovatnoća klase *i*, a *C* ukupan broj klasa.

Relativna entropija meri smanjenje entropije nakon što se skup podataka podeli na bazi atributa.

$$IG = Entropija(roditelj) - \sum_{i=1}^{n} \frac{|dete_i|}{|roditelj|} Entropija(dete_i) \#(3)$$

Informacioni dobitak (IG) meri koliko se „neizvesnost" smanjuje kada podelimo skup podataka na podskupove na osnovu određene karakteristike (atributa). Entropija (roditelj) meri stepen entropije pre podele, dok je sledeca komponenta u jednačini 3 zapravo prosečna entropija podskupova nakon podele.

$$Entropija = -\sum_{i=1}^{C} p_i \log_2(p_i) \#(4)$$

U formuli 4 definisana je **entropija**, koja zapravo meri stepen neizvesnosti u raspodeli podataka. Konkretno time se kvantifikuje stepen pomešanosti podataka (visoka entropija znači da je heterogenost podatkaa visoka). Ključni delovi su:

- *C* - Broj klasa (npr. pozitivno i negativno kod binarne klasifikacije).
- $p_{_i}$ - Verovatnoća pripadnosti uzorka klasi *i*.

DT modeli su jednostavni za razumevanje i interpretiranje, što je zapravo jedan od faktora zašto su često primenjeni u praksi. Pored toga, stabla odlučivanja mogu da obuhvate nelinearne odnose između karakteristika i ciljne promenljive, što je korisno kada je odnos između promenljivih složen. Takođe, primena RF modela zahteva manje pripreme podataka u poređenju sa drugim metodama (npr. ne zahtevaju normalizaciju ili skaliranje podataka).

Sa druge strane postoji nekoliko očiglednih nedostataka modela DT, kao što su *nestabilnost, pristrasnost i preterano prilagođavanje*. Konkretno male promene u podacima mogu dovesti do potpuno drugačije strukture stabla, čineći stabla odluka manje stabilnim u poređenju sa drugim modelima. Pored toga DT modeli su skloni pristrasnosti prema klasama koje su dominantne u uzorku. Tako model može zanemariti manje zastupljenu, ali važnu klasu u uzorku. I na kraju, čest problem DT modela je preterano prilagođavanje, što znači da

neki specifični podaci (ili šum u podacima) u trening uzorku mogu biti zapamćeni i mogu uticati na loše performanse prilikom primene na stvarnim podacima.

2.3 Nasumične šume (RF) pripada ansambl metodama mašinskog učenja koja funkcioniše tako što kombinuje više stabala odlučivanja, a cilj je da se poboljša tačnost i poveća stabilnost odluka[13].

Model RF funkcioniše tako što gradi nekoliko stabala odlučivanja koristeći različite podskupove podataka i karakteristika. Svako stablo odlučivanja u celokupnoj strukturi je izgrađeno na uzorku iz originalnih podataka, a pri svakom razdvajanju u stablu, slučajni podskup karakteristika se razmatra kao potencijalni kriterijum za razdvajanje. Ova nasumičnost pomaže u smanjenju varijanse modela, čineći ga robusnijim i manje sklonim preteranom prilagođavanju[14].

Konkretno proces kreiranja RF funkcioniše tako da se iz orginalnog skupa podataka D metodom ponovnog izvlačenja uzoraka generiše n novih uzoraka. $D_b = \{(x_i, y_i) \mid (x_i, y_i)D, \text{za } i = 1,2, \dots, n\}$. Na svakom od čvorova stabla, jedna tj. najbolja karakteristika iz podskupa m iz ukupnog broja t karakteristika je odabrana kao kriterijum za selekciju. Nakon toga, kao i kod DT, Gini indeks i Entropija se koriste za merenje kvaliteta podele. Takođe može se koristit IG kako bi se najbolje izmerila podela putem merenja smanjenja entropije.

Postoji nekoliko ključnih prednosti zbog kojih je RF model izuzetno široko rasprostranjen i često primenjivan za rešavanje različitih problema. U poređenju sa DT modelom, RF daje daleko preciznije rezultate jer se agregiranjem i nalaženjem proseka rezultata velikog broja stabala dobijaju tačnija predviđanja.

Međutim, postoji i nekoliko nedostataka ovog metoda. Kombinovanjem rezultata više stabala odlučivanja, algoritam postaje složeniji i manje interpretabilan u poređenju sa pojedinačnim stablom odlučivanja, što dalje utiče na povećanje vremena analize i zahteva snage računara[15].

2.4 Duboke neuronske mreže (DNN) su vrsta mašinskog učenja nastala inspirisana strukturom i načinom funkcionisanja ljudskog mozga. Sastoji se od više međusobno povezanih slojeva neurona koji zajedno obrađuju ulazne podatke i proizvode izlaze rezultate. Svaki sloj transformiše ulazne podatke u

[13] Breiman, L., (2001), p.7-14
[14] Lessmann, S., Baesens, B., Seow, H. V., i Thomas, L. C. (2015), p. 107
[15] Liu, Y., Huang, F., Ma, L., & Shi, J. (2024), p. 2-6

sve apstraktnije reprezentacije, omogućavajući mreži da nauči složene obrasce i odnose[16].

DNN se sastoji od ulaznog sloja, nekoliko skrivenih slojeva i izlaznog sloja. Svaki sloj sadrži skup neurona (čvorova) koji obrađuju podatke. Veze između neurona imaju posebne težine (pondere) koji se prilagođavaju tokom treninga kako bi se kontinuirano smanjivala greška između predviđanja i stvarnih ishoda[17]. Na konkretnom primeru kreditnog skoringa, ulazne karakteristike su kreditna istorija, odnos duga i prihoda, status zaposlenja i slično. Nakon toga aktiviraju se skriveni slojevi koji transformišu ulazne podatke u pomenute apstraktne reprezentacije. Svaki neuron u skrivenom sloju obavlja proračun koji uključuje ponderisani zbir ulaza praćenih funkcijom aktivacije[18].

U jednačini ispod *h* predstavlja autput neurona *i* u sloju *l*

$$h_i^{(l)} = f(\sum_j w_{ij}^{(l-1)} h_j^{(l-1)} + b_i^{(l)})\#(5)$$

gde su: wij(l-1) je težina veze između neurona j iz sloja l-1 i neurona i iz sloja l ; hj(l-1) rezultat neurona j iz prethodnog sloja; bi(l) pristrasnost neurona; f funkcija aktivacije transformiše linearnu kombinaciju ulaza (elementi unutar zagrade) u izlaz.

Funkcije aktivacije određuju rezultat neurona i uvode nelinearnost u mrežu, čime joj se daje mogućnost da nauči složene procese. Sigmoidna funkcija se često koristi, a njena rešenja se kreću u intervalu od 0 do 1, što je izuzetno pogodno za probleme, poput kreditnog skoringa, gde se rezultat interpretira kao verovatnoća[19]. Ispravljač ReLU je popularna za skrivene slojeve u DNN-ovima zbog svoje jednostavnosti i mogućnosti da mreža uči brže i bolje radi sa velikim skupovima podataka[20].

DNN su pogodne sa modeliranje složenih, nelinearnih odnosa i interakcija između karakteristika, što dovodi do visoke preciznosti predviđanja. Pored toga, DNN mogu da se koriste na različitim tipovima podataka, kako struktuiranim tako i na nestruktuiranim podacima, poput slika, govora, video materijala i slično[21].

Ipak evidentno je postojanje nekoliko ključnih nedostataka, prvenstveno to da je proces modeliranja DNN-a manje transparentan i dosta komplikovan u

[16] Hayashi, Y. (2022), p. 4-12
[17] LeCun, Y., Bengio, Y. and Hinton, G. (2015), p. 5-9
[18] Chang, A.-H.,et al. (2022), p. 310
[19] Shi, S. T., (2022), p. 14331
[20] Mestiri, S., (2024), p. 243
[21] Chang, A.-H.,et al. (2022), p.312

poređenju sa ostalim modelima, što otežava tumačenje rezultata i njihovu validaciju, pogotovo ako govorimo u kontekstu regulatornog tela. Pored toga sam proces izračunavanja može vremenski biti izuzetno zahtevan. Na kraju, DNN, slično DT i RF modelima, često pate od preteranog prilagođavanja ukoliko su podaci za obuku mali ili sadrže neki šum[22].

3. Merenje performansi modela

Metrika koja se koristi za analizu i rangiranje performansi modela kreditnog skoringa je od ključnog značaja kako bi se procenilo koji model najbolje funkcioniše, odnosno koji model najtačnije procenjuje dužnike koji neće moći da otplate dugovanja. Modeli kreditnog skoringa se obično procenjuju korišćenjem različitih indikatora performansi koje pružaju uvid u njihovu tačnost i diskriminatornu sposobnost. Ovi indikatori pomažu finansijskim institucijama da donesu odluku o odabiru modela.

Tačnost se računa kao odnos tačno predviđenih i ukupnog broja predviđanja. Ovaj indikator pruža osnovnu meru ukupne ispravnosti modela. U jednačini ispod, TP (*engl.* True Positive) predstavlja predviđanje u kojim model tačno predvidi slučajeve koji zaista nisu otplatili kredit. TN (*engl.* True Negative) je kada model tačno predividi one koji uredno plaćaju svoj kredit. FP (*engl.* False Positive) su slučajevi gde je model predvideo bankrot u slučajevima gde se to nije deslio. I na kraju, FN (*engl.* False Negative) su pogrešna predviđanja da neće doći do bankrota u slučajevima kada dužnik zapravo doživi bankrot[23].

$$\text{Tačnost} = \frac{TP + TN}{TP + TN + FP + FN} \#(6)$$

Iako je tačnost jednostavna metrika, ona može biti problematična u skupovima podataka gde broj klijenata koji su bankrotirali i onih koji nisu, nije izbalansiran. Na primer, ako skup podataka ima 95% slučajeva koji nisu bankrotirali, model koji uvek predviđa da nema bankrota će postići tačnost od 95%, iako ne uspe da identifikuje nijedan slučaj klijenata koji je bankrotirao.

Preciznost meri tačnost pozitivnih predviđanja napravljenih od strane modela. Označava koliko je od predviđenih slučajeva bankrota zaista bankrotiralo. Preciznost je ključna kada su troškovi pogrešnih pozitivnih rezultata visoki, odnosno slučajeva u kojima je odbijen kreditni zahtev tražiocima koji su sposobni za vraćanje kredita[24].

[22] Edunjobi, T. E., et al. (2024), p. 101
[23] Zou, Y., & Gao, C. (2022), p.10
[24] Mestiri, S. (2024), p. 243

$$Preciznost = \frac{TP}{TP+FP} \#(7)$$

Opoziv meri sposobnost modela da ispravno identifikuje sve slučajeve kada dođe do bankrota i ukazuje na proporciju realizovanih bankrota koji su predviđeni kao problematični platioci. Opoziv je od suštinskog značaja kada su troškovi pogrešnih negativnih odluka visoki, kao što je odobravanje kredita visokorizičnom dužniku za koji postoji verovatnoća da neće platiti kredit[25].

$$Opoziv = \frac{TP}{TP + FN} \#(8)$$

F1 rezultat je sredina *Preciznosti* i *Opoziva*, čime se obezbeđuje ravnoteža između ove dve metrike. Time F1 biva korisni pokazatelj kada postoji neujednačena distribucija klasa. F1 rezultat je posebno bitan u situacijama kad su pogrešno pozitivna i negativna predviđana dominantna. Što dalje znači da je izuzetno korisna mera prilikom analize gde je neophodno postići kompromis između *Preciznosti* i *Opoziva*.

$$F1 = 2 \times \frac{\text{Preciznost} \times \text{Opoziv}}{\text{Preciznost} + \text{Opoziv}} \#(9)$$

AUC - ROC (*engl.* Area Under Curve **-** Reciever Operating Characteristics) koeficijent predstavlja zapravo oblast ispod ROC krive. Koristeći primer modela kreditnog skoringa predstavlja procenu sposobnosti modela da pravi razliku između slučajeva koji su bankrotirali i onih koji nisu. Vrednost ovog koeficijenta se kreće od 0 do 1, gde 0,5 označava da nema diskriminatorne moći, odnosno da je predviđanje bazirano na slučajnom hodu, a 1 predstavlja slučaj u kom model savršeno predviđa bankrot, odnosno regularno vraćanje kredita[26]. U jednačini ispod, *TPR* (*engl.* True Positive Rate) predstavlja tačno klasifikovane pozitivne slučajeve, odnosno procenat klijenata koji zaista neće vratiti kredit koje je model tačno predvideo.

$$AUC - ROC = \int_0^1 \text{TPR}(x) \#(10)$$

Gini koeficijent je mera nejednakosti ili koncentracije, koja se obično koristi za procenu diskriminatorne moći modela kreditnog skoringa i usko je povezan je sa AUC-ROC

$$Gini = 2 \times \text{AUC-ROC} - 1 \#(11)$$

Gini koeficijent se kreće od -1 do 1, pri čemu 0 označava da nema diskriminacije, a 1 ukazuje na savršenu diskriminaciju. Veći Gini koeficijent

[25] Chang, A.-H.,et al. (2022), p.313
[26] Mestiri, S. (2024), p. 243

ukazuje na bolje performanse modela u razlikovanju između slučajeva koji su bankrotirali i onih koji redovno ispunjavaju svoje obaveze[27].

4. Baza podataka

Za ovo istraživanje koristićemo skup podataka iz „Kaggle“ baze podataka „Give Me Some Credit“, koji je javno dostupan. Ovaj skup podataka je široko priznat resurs za istraživanje kreditnog skora i sadrži niz karakteristika koje su relevantne za procenu kreditnog rizika. Skup podataka pruža istorijske podatke o zajmoprimcima, uključujući njihovu finansijsku metriku i kreditno ponašanje, koji se mogu koristiti za izgradnju prediktivnih modela za određivanje verovatnoće da klijent neće vratiti kredit. Skup podataka „Give Me Some Credit“ se sastoji od 150.000 redova i 11 kolona. Svaki red odgovara pojedinačnom dužniku, a svaka kolona predstavlja različite karakteristike tj. promenljive povezane sa demografskim karakteristikama i finansijskim profilom dužnika. Primarni cilj je da se predvidi da li će dužnik tokom perioda kreditiranja upasti u finansijske probleme koji će uzrokovati kritično kašnjenje u izmirenju obaveza.

Tabela 1. Pregled promenljivih u bazi podataka

Promenljiva	**Opis**	**Tip**
Kritično kašnjenje	Kašnjenje duže od 90 dana u poslednje dve godine	Broj
Ostala neosigurana dugovanja	Odnos ostalih ukupnih dugovanja u odnosu na kreditni limit	Procenat
Starost	Starost dužnika izražena u godinama	Broj
Kasnio do 60 dana	Broj kašnjenja do 60 dana u poslednje dve godine	Broj
Odnos duga	Ukupan dug podeljen sa mesečnim prihodom	Procenat
Godišnja zarada	Godišnja zarada	Broj
Broj kredita i kreditnih linija	Broj kredita i kreditnih linija	Broj
Kasnio 90 i duže	Broj kašnjenja do 90 dana u poslednje dve godine	Broj
Broj hipotekarnih kredita	Broj hipotekarnih kredita	Broj
Kasnio između od 60 do 90 dana	Broj dana kašnjenja od 60 do 90 dana u poslednje dve godine	Broj
Broj članova porodice	Broj članova porodice koji zavise od tražioca kredita	Broj

[27] Chang, A.-H., et al (2022) p.310-312

Razumevanje distribucije i karakteristika podataka je ključna za izgradnju efektivnih modela kreditnog ocenjivanja. Ispod je sažetak opisne statistike za svaku funkciju u skupu podataka.

Tabela 2. Deskriptivna statisitka

Promenljive	**Broj opservacija**	**Srednja vrednost**	**Standardna devijacija**	**Minimum**	**Maksimum**
Kritično kašnjenje	150000.00	0.07	0.25	0.00	1.00
Ostala neosigurana dugovanja	150000.00	6.05	249.76	0.00	50708.00
Starost	150000.00	52.30	14.77	0.00	109.00
Kasnio do 60 dana	150000.00	0.42	4.19	0.00	98.00
Odnos duga	150000.00	353.01	2037.82	0.00	329664.00
Godišnja zarada	120269.00	6670.22	14384.67	0.00	3008750.00
Broj kredita i kreditnih linija	150000.00	8.45	5.15	0.00	58.00
Kasnio 90 i duže	150000.00	0.27	4.17	0.00	98.00
Broj hipotekarnih kredita	150000.00	1.02	1.13	0.00	54.00
Kasnio od 60 do 90 dana	150000.00	0.24	4.16	0.00	98.00
Broj članova porodice	146076.00	0.76	1.12	0.00	20.00

Izvor: Baza „Give me some credit” iz Kaggle

Posmatrajući promenljive i njihove raspodele u tabeli 2 evidentno je nekoliko nedostataka, koji ukazuju na neophodnost dalje transformacije, kako bi rezultati bili smisleniji i relevantniji:

- Ostala neosigurana dugovanja, mesečni dohodak i odnos duga pokazuju izuzetnu zakrivljenost zbog prisustva ekstremnih vrednosti. Stoga je u cilju smanjenja zakrivljenosti, neophodno limitirati ekstremne vrednosti korišćenjem neke od adekvatnih tehnika. U ovom radu konkretno je korišćena „winzorizacija" (*engl*. Winsorizing).
- Mesečni dohodak i broj članova domaćinstva imaju veliki broj nedostajućih vrednosti. Ovo je pogotovo problem kod promenljive

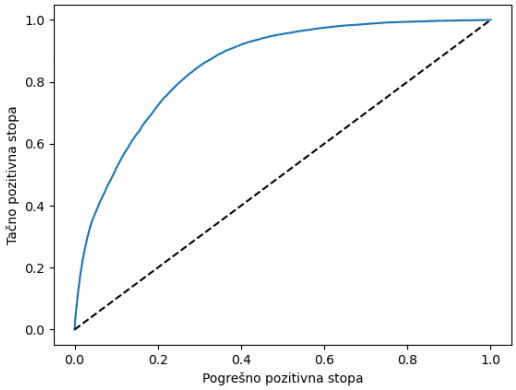
mesečnog dohotka jer je ona od esencijalnog značaja za analizu, tako da ćemo primeniti tehniku imputacije srednje vrednosti uzorka.

- Na kraju zavisna promenljiva, (kašnjenje duže od 90 dana) pokazuje izuzetno neizbalansiranu strukturu. Samo 6,68% tražilaca kredita ima neizmirivane obaveze. Ova neravnoteža će svakako biti u fokusu tokom istraživanja i biće primenjena SMOTE (*engl.* Synthetic Minority Oversampling Technique) tehnika kako bi observacije koje su manje zastupljene u uzorku bile adekvatno tretirane[28].

5. Empirijski rezultati. Nakon inicijalne pripreme i prilagođavanja podataka u svrhu dalje empirijske analize, u narednom koraku, prethodno objašnjena četiri modela, logistička regresija (LR), nasumične šume (RF), stabla odlučivanja (DT) i duboke neuronske mreže (DNN) biće analizirana i međusobno upoređena po principu efikasnosti. Za analizu će biti korišćen programski jezik „Python“ i to konkretno „PyCharm“ programsko okruženje koje nudi optimalne uslove za efikasnu analizu. Osnova, odnosno metrike za evaluaciju biće tačnost, F1-rezultat, oblast ispod krive (AUC), opoziv, preciznost i Gini koeficijent, čiji su teorijski okrviri takođe objašnjeni u prethodnom delu.

Logistička regresija daje pričino solidne rezultate ukupno gledano. Ako pak detaljnije analiziramo AUC krivu (plava linija) u poređenju sa slučajnom kretnjom (isprekidana crna linija) daleko je superiornija (slika 1). Međutim, činjenica da se bavi isključivo jednostavnim vezama ograničava sposobnost da opiše složene odnose u skupu podataka.

Slika 1. AUC kriva logističke regresije

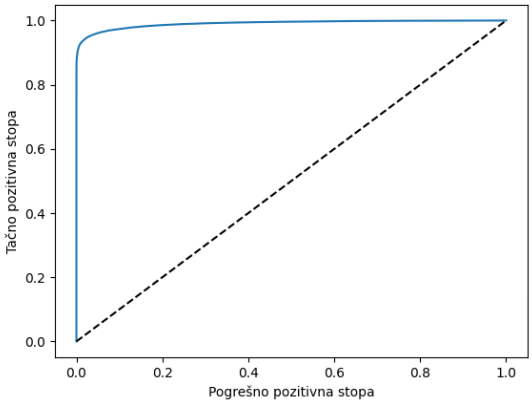


Nasumične šume (RF) pokazuju značajno veću prediktivnu moć u poređenju sa logističkom regresijom, a po nekim parametrima (AUC i Gini) čak

[28] Niu, A., (2020), p. 4

i od stabla odlučivanja (tabela 3). Izuzetno visoka vrednost AUC (slika 2) i Gini koeficijenta pokazuje efikasnost, odnosno naglašava robustnost modela, pogotovo kada se radi o modeliranju nelinearnim obrascima i interakcijama između promenljivih.

Slika 2. AUC kriva modela nasumičnih šuma



Model stabla odlučivanja nadmašuje, logističku regresiju i nasumične šume, po preciznosti, tačnosti, F1 i opozivu (tabela 3). Međutim ovaj model ima tendenciju preteranog prilagođavanja, što može smanjiti njegovu efikasnost u manjim skupovima podataka. Ipak treba dodati da se ovaj efekat može elegantno smanjiti (možda i potpuno eliminisati) koristeći tehniku pojačavanja (*engl.* Boosting).

Tabela 3. Empirijski rezultati

Model	**Tačnost**	**F1**	**AUC**	**Opoziv**	**Preciznost**	**Gini**
Logistička regresija	76.94%	76.91%	0.8528	76.82%	76.99%	0.7055
Nasumične šume	88.95%	88.88%	0.9553	88.25%	89.51%	0.9106
Drvo odlučivanja	93.58%	93.61%	0.9358	94.03%	93.19%	0.8716
DNN	**96.11%**	**96.03%**	**0.9906**	**94.06%**	**98.09%**	**0.9813**

Duboka neuronska mreža pokazuje najbolje performanse po svim metrikama. Ovi rezultati naglašavaju sposobnost DNN-a da uhvati složene, nelinearne odnose unutar podataka, iako po cenu povećane složenosti i smanjene transparentnosi. Podešavanje hiperparametara, dizajn arhitekture i paž-

ljivo rukovanje prekomernim prilagođavanjem su od suštinskog značaja da bi se u potpunosti iskoristio potencijal DNN-a.

Ukupno gledano možemo izvući tri ključna zaključka analize:

1. Tradicionalni modeli, koji su danas najzastupljeniji u praksi, predstavljaju dobru osnovu za analizu kreditnog skoringa, pogotovo zbog svoje jednostavnosti tumačenja. Međutim, upravo to i jeste značajan ograničavajući faktor, jer nije moguće predstaviti kompleksne veze.
2. Napredni modeli pokazuju jasnu dominaciju u odnosu na logističku regresiju prema skoro svim metrikama. Pogotovo je DNN model potvrdio svoju dominaciju u odnosu na sve ostale, što definitivno ukazuje na opravdanost uključivanja istih u analize ukoliko je cilj poboljšanje predviđanja i tačnosti (tabela 3).
3. Ipak, posmatrajući obe dimenzije, jednostavnost i transparentnost sa jedne strane, kao i tačnost i preciznost sa druge, najlogičnije rešenje je hibridna verzija, odnosno integracija i korišćenje prednosti obe strane.

6. Zaključak

Prethodno izneto poređenje modela logističke regresije, stabla odlučivanja, nasumične šume i dubokog učenja, korišćenjem neke od metrika evaluacije (tačnost, preciznost, gini, F1, opoziv i AUC) predstavlja važnu polaznu osnovu za promenu dominantne pozicije jednostavnih modela kreditnog skoringa u bankraskoj praksi. Obzirom da je logistička regresija pokazala inferiornost u odnosu na sve ostale modele po skoro svim kriterijumima lako možemo doći do zaključka da i u praksi napredni modeli mogu povećati efikasnost kreditnog skoringa, što dalje ima praktične implikacije na poboljšanje upravljanja rizicima u bankama i drugim finansijskim institucijama.

Ipak u svrhu šire praktične primene, ovoj studiji treba dodati još nekoliko dimenzija čime bi rezultati i zaključci bili robusniji. *Prva* je proširenje baze podataka koja će biti korišćena za treniranje i testiranje modela, po više kriterijuma. Tu se misli na uključivanje podataka iz različitih država čime bi se dobila geografski diverzifikovan pogled. Nakon toga bi trebalo uključiti nestruktuirane podatke u analizu kako bi se pokušali uhvatiti uticaji koji nisu identifikovani do sada zbog nedostatka podataka i mogućnosti da se analiziraju[29].

[29] Talaat, F. M., et al.,(2023), p. 4-7

Druga važna dimenzija i pravac proširenja jeste uključivanje modela generativne veštačke inteligencije (GAN, VAE, LSTM i slični) u analizu čime bi se značajno doprinelo „obogaćivanju" podataka sa dimenzijama koje su obično u realnom uzorku ređe zastupljene.

Poslednja, *treća* važna dimenzija jeste interpretabilnost i transparentnost samih modela. Iako je ključna prednost statističkih modela upravo jednostavnost i razumljivost, upravo to je važan nedostatak naprednih modela mašinskog i dubokog učenja. Razvijanje objašnjivih metoda veštačke inteligencije, kao što su SHAP ili LIME, može da odgovori na zabrinutost oko interpretabilnosti složenih modela i da doprinese većoj transparentnosti[30].

Na kraju, treba reći da je ovo istraživanje na pravi način istaklo važnost inovativnih i modernih modela koji su postali dostupni napretkom tehnologije, i ukazuje da njihova primena može doprineti kvalitetnijem upravljanju rizika[31]. Stoga bi trebalo raditi na ostalim komponentama, pored istraživanja, kao što je regulativa, kako bi se ovi modeli mogli primenjivati u praksi.

Literatura:

Abel, A., Bernanke, B. i Croushore, D. (2007), Macroeconomics, Addison–Wesley. ISBN 978-0-321-41554-7.

Addy, W., Ajayi-Nifise, A., Bello, B., Tula, S., Odeyemi, O., Falaiye, T. (2024). AI in credit scoring: A comprehensive review of models and predictive analytics. 18. 118-129. https://doi.org/10.30574/gjeta.2024.18.2.0029

Agarwal, S., Alok, S., Ghosh, P., i Gupta, S. (2023), "Role of Big Data and Machine Learning in Fintech", Global Finance Journal.

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., i Vanthienen, J. (2003), "Benchmarking state-of-the-art classification algorithms for credit scoring", Journal of the Operational Research Society, Vol. 54(6), Pp. 627-635. https://doi.org/10.1057/palgrave.jors.2601545

Barddal, J. P., Loezer, L., Enembreck, F., i Lanzuolo, R. (2020), "Lessons learned from data stream classification applied to credit scoring", Expert Systems With Applications. https://doi.org/10.1016/j.eswa.2020.113899

Breiman, L., Friedman, J.H., Olshen, R.A., i Stone, C.J. (1984). Classification and Regression Trees. Biometrics, 40, 874. https://doi.org/ 10.2307/2530946

Breiman, L., (2001),"Random Forests". Machine Learning 45, 5-32. https://doi.org/ 10.1023/A:1010933404324

[30] Misheva, H., (2021). P.2-11

[31] Lukić, V. i Popović, S. (2020), p. 56 – 62

Chang, A.-H., Yang, L.-K., Tsaih, R.-H., i Lin, S.-K. (2022), "Machine learning and artificial neural networks to construct P2P lending credit-scoring model: A case using Lending Club data", Quantitative Finance and Economics, Vol. 6(2), Pp. 303-325, https://doi.org/10.3934/QFE.2022013

Dumitrescu, E. i Hué, Sullivan & Hurlin, Christophe & Tokpavi, Sessi, (2022) "Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects", European Journal of Operational Research, Elsevier, vol. 297(3), strane 1178-1192. https://doi.org/10.1016/j.ejor.2021.06.053

Edunjobi, T. E., i Odejide, O. A. (2024), "Theoretical frameworks in AI for credit risk assessment: Towards banking efficiency and accuracy", International Journal of Scientific Research Updates, Vol. 7(1), Pp. 092-102, https://doi.org/10.53430/ijsru.2024.7.1.0030

Hayashi, Y. (2022), "Emerging Trends in Deep Learning for Credit Scoring: A Review", Electronics, Vol. 11(3181), https://doi.org/10.3390/electronics11193181.

LeCun, Y., Bengio, Y. i Hinton, G. (2015), "Deep Learning. Nature", 521, 436-444. https://doi.org/10.1038/nature14539

Lessmann, S., Baesens, B., Seow, H. V., i Thomas, L. C. (2015), "Benchmarking state-of-the-art classification algorithms for credit scoring: A ten-year update", Journal of the Operational Research Society, Vol. 66(1), Pp. 101-113. https://doi.org/10.1057/jors.2014.20

Lukić, V. i Popović, S., (2020.), "Sistematska Kriza Problematičnih Kredita I Ekonomska Politika - Analiza Iskustava Nakon Globalne Finansijske Krize (Systemic Crisis Of Nonperforming Loans And Economic Policy - Analysis Of Experiences", Ekonomske ideje i praksa, broj37, strane 53-78, Jun.

Liu, Y., Huang, F., Ma, L., i Shi, J. (2024), "Credit scoring prediction leveraging interpretable ensemble learning", Journal of Forecasting, https://doi.org/10.1002/for.3033

Niu, A., Cai, B. i Cai, S. (2020). "Big Data Analytics for Complex Credit Risk Assessment of Network Lending Based on SMOTE Algorithm, Complexity". 2020.1-9. https://doi.org/10.1155/2020/8563030

Mestiri, S. (2024), "Credit scoring using machine learning and deep Learning-Based models", Data Science in Finance and Economics, https://doi.org/10.3934/DSFE.2024009

Miloš Božović, (2021.) "Sistemska Komponenta Kreditnog Rizika: Metod Kopula (Systemic Component Of Credit Risk: A Copula-Based Method)," Ekonomske ideje i praksa, broj 41, strane 1-13, Jun.

Misheva, H., Osterrieder, B., Hirsa, J., Kulkarni, A., Lin, S. (2021). "Explainable AI in Credit Risk Management", 10.48550/arXiv.2103.00949.

Shi, S., Tse, R., Luo, W., D'Addona, S., i Pau, G. (2022), "Machine learning-driven credit risk: a systemic review", Neural Computing and Applications, https://doi.org/10.1007/s00521-022-07472-2

Talaat, F. M., Aljadani, A., Badawy, M., i Elhosseini, M. (2023), "Toward interpretable credit scoring: integrating explainable artificial intelligence with deep learning for credit card default prediction", Neural Computing and Applications, https://doi.org/10.1007/s00521-023-09232-2

Thomas, L. C., Edelman, D. B., i Crook, J. N. (2004). "Readings in credit scoring: foundations, developments, and aims". Oxford University Press. https://doi.org/10.1093/oso/9780198527978.001.0001

Primljen (Received): 15.1.2025.
Prihvaćen (Accepted): 8.2.2025.
Pre štampe (Online First) 21.2.2025.