

Bayesov pristup logističkoj regresiji za predviđanje stečaja trgovačkih društava u uvjetima neizvjesnosti

Sažetak

Ovaj članak pruža odgovore na česta metodološka pitanja u istraživanjima koja se bave predviđanjem stečaja trgovačkih društava, kao što su odabir homogenog slučajnog uzorka, selekcija financijskih i nefinancijskih pokazatelja relevantnih za predviđanje stečaja, odabir optimalne granične vjerojatnosti za klasifikaciju te specifikacija prior distribucija za primjenu Bayesove logističke regresije, i stoga predstavlja dodanu vrijednost postojećoj literaturi. S obzirom na potrebu za predviđanjem stečaja kao najučinkovitijim načinom njegove prevencije, posebice u kriznim vremenima, nameće se zahtjev za unaprijeđenjem prognostičkih modela. U tom je kontekstu istraženo, ne samo koji su pokazatelji relevantni za predviđanje stečaja u djelatnosti veleprodaje i maloprodaje, već pridonosi li Bayesov pristup prediktivnoj sposobnosti logističke regresije u razdoblju neizvjesnosti, tj. u razdoblju COVID pandemije u kojem je istoimena metoda i primijenjena. Prikladnost Cauchyjevog priora za konstantni član je potvrđena, dok je za preostale koeficijente Studentov t-prior sa standardnom devijacijom od 2,5 na ljestvici logaritamskih omjera šansi, primjenjiviji. Obrtaj kratkotrajne imovine je pokazatelj koji najbolje diskriminira stečajna od ne stečajnih trgovačkih društava, iako je najviše neizvjestan. Unatoč tome, EBIT marža pokazuje najmanju diskriminacijsku moć, iako je najizvjesnija varijabla, dok su zaduženost i tekuća likvidnost umjereno neizvjesne varijable. Prediktivna sposobnost Bayesove logističke regresije od 90,22% u 2021. godini potvrđena je križnom validacijom. Pritom je točno klasificirano 99,61% stečajnih trgovačkih društava, ali s nižom graničnom vjerojatnosti u usporedbi sa standardnom logističkom regresijom. Adekvatno specificirani slabo informativni priori, koji su ključni za posteriornu analizu parametara, glavni su doprinos istraživanja, a posljedično i određivanje razine neizvjesnosti pojedinačnih indikatora stečaja.

Ključne riječi: Bayesova logistička regresija, predviđanje stečaja, financijski i nefinancijski pokazatelji, trgovina na veliko i malo, prior distribucije parametara

^a J. Arnerić, dr. sc., redoviti profesor, Sveučilište u Zagrebu, Ekonomski fakultet (e-mail: jarneric@efzg.hr).

^b M. Moćan, student diplomskog sveučilišnog studija Poslovne ekonomije na Ekonomskom fakultetu Sveučilišta u Zagrebu (e-mail: mocan.matteo@gmail.com). Rad je primljen u uredništvo 16.10.2023. godine, a prihvaćen je za objavu 11.04.2024. godine.

Ovaj članak proizašao je iz obranjenog diplomskog rada na integriranom preddiplomskom i diplomskom sveučilišnom studiju Poslovne ekonomije Ekonomskog fakulteta Sveučilišta u Zagrebu.

1. UVOD

Kada dužnik postane dugotrajno nesposoban podmirivati svoje dospjele financijske obveze, bilo da se radi o ugovornim ili zakonskim obvezama, te je prisiljen prodati imovinu kako bi pokrio troškove stečajnog postupka, on se nalazi u stečaju. Pojam stečaj potječe iz njemačkog pojma "konkurs", što doslovce znači "strka vjerovnika". Također, koriste se i pojmovi poput "bankrot" (engl. *bankruptcy*), "neuspjeh" (engl. *failure*) ili "financijska poteškoća" (engl. *financial distress*). Stečaj se događa kada se ispunjava zakonom propisan razlog za stečaj, koji može uključivati i "smrt" pravne osobe, a uzroci stečaju mogu biti interni i eksterni. Interni uzroci uključuju loše upravljanje tvrtkom, dok su vanjski uzroci opći te obuhvaćaju pravne, ekonomske, socijalne, ekološke i tehnološke čimbenike. Ooghe i De Prijcker (2008) navode manjak menadžerskog znanja i nedostatak financijskih sredstava kao temeljne uzroke stečaja. Ostali uzroci uključuju recesiju u industriji, nedostatak povjerenja potrošača, promjene zakona, inertnost i nedostatak motivacije vodstva te jaka konkurencija (Sajter, 2014). Novak i Sajter (2007) istraživali su razloge stečaja u Hrvatskoj te su identificirali nekoliko ključnih čimbenika, kao što su nedovoljno razvijeno tržište, zastarjelu tehnologiju, posljedice rata, loš menadžment, prijevaru kreditora, loše zakonodavstvo, nedostatak osobne odgovornosti za dugove, slabu financijsku disciplinu, loše planiranje te problematične međuljudske odnose. U drugim zemljama, primjerice u Engleskoj, prezaduženost, gubitak tržišta zbog nedostatka prodaje i jake konkurencije, te problemi s menadžmentom, uključujući loše planiranje i prevelik optimizam, neplaćanje poreza, česti su uzroci stečaja. U Njemačkoj su se kao uzroci stečaja pokazali neočekivani problemi s projektima, loš menadžment otplate dugova, poslovanje bez kontrolinga te slaba komunikacija.

Stečajni postupak uključuje više dionika, ne samo vjerovnike i vlasnike, već i zaposlenike, na koje ima izravan ili neizravan utjecaj. Vjerovnicima, poput banaka ili dobavljača s odgodom plaćanja, prijeti gubitak posuđenih sredstava, a u slučaju njihovog povrata prate ih pravni troškovi, kašnjenja u plaćanju i poremećaji u novčanom tijeku. Ako postoji visoka ovisnost o drugim poduzećima, može doći do lančane reak-

cije, gdje jedan stečaj može izazvati niz drugih. Vjerovnici se prvo namiruju jer su svoje dugove osigurali i po zakonu su prioritetni. Preostali novac iz likvidacijske mase isplaćuje se povlaštenim dioničarima, a zatim držateljima običnih dionica. Zaposlenici su i vjerovnici i dionici u stečajnom postupku jer tvrtkauguje plaće ili druge naknade vezane uz rad, a istovremeno im prijeti otkaz. Glavna svrha sprječavanja stečaja je zaštita zaposlenika, investitora, vjerovnika, poslovnih partnera pa čak i kupaca, iako ostaje pitanje žele li vođe poduzeća sami izbjeći stečaj i trebaju li poboljšati kadar, financije, marketing, istraživanje i razvoj te strategiju da bi stečaj prevenirali.

Većina autora u dosadašnjim istraživanjima koristi mali uzorak poduzeća ili uzorak koji je gotovo jednak broju varijabli, što dovodi do neefikasnih i nesignifikantnih procjena parametara zbog malog broja stupnjeva slobode. Osim toga, mnogi u uzorak izabiru jednak broj uspješnih i neuspješnih poduzeća, prilagođavajući se razini praga vjerojatnosti od 0,5 što nije ispravno jer takvi uzorci nisu reprezentativni s obzirom na broj uspješnih i neuspješnih poduzeća u populaciji. Pristup "uparenih" uzoraka je pogrešan i pruža prividno dobru prediktivnu točnost modela logističke regresije. Nedostatak je i što financijski pokazatelji često nisu normalno distribuirani, dočim je pretpostavka normalne distribucije ključna prilikom procjene parametara metodom najveće vjerodostojnosti (engl. *maximum likelihood*). S druge strane, logistička regresija, kao alternativa diskriminacijskoj analizi, razvijena je kako bi se omogućila analiza podataka u uvjetima kada je pretpostavka normalnosti narušena. Međutim, i logistička regresija i diskriminacijska analiza pretpostavljaju da su parametri fiksni, dok su varijable slučajne, što ih čini frekventističkim pristupima.

Za razliku od frekventista, Bayesovci pretpostavljaju da su parametri slučajni, dok su varijable fiksne. Bayesov pristup zahtijeva prethodno poznavanje distribucije vjerojatnosti svakog parametra, a potom se izvode njihove posteriorne distribucije i posteriorni intervali neizvjesnosti. Zapravo, Bayesov pristup omogućuje određivanje cijele distribucije omjera šansi (engl. *odds ratios*), što je posebno korisno u uvjetima neizvjesnosti, poput onih tijekom COVID-19 pandemije.

Stoga je glavni istraživački cilj utvrditi doprinos Bayesovog pristupa logističkoj regresiji u predviđanju stečaja tijekom COVID-19 pandemije, koja je sigurno utjecala na stečajeve trgovačkih društava u sektoru trgovine u Republici Hrvatskoj (Keglevi i sur., 2022). Domaća literatura do sada nije analizirala predviđanje stečaja u sektoru trgovine primjenom Bayesove logističke regresije (BLR) u uvjetima neizvjesnosti. Iako postoje radovi u Hrvatskoj koji su istraživali različite sektore, uključujući trgovinu, pružaju različite rezultate zbog varijacija u opsegu podataka. U tim radovima rijetko je elaboriran način slučajnog uzorkovanja poduzeća za analizu, a autori se rijetko osvrću na ograničenja primijenjenog pristupa, te ne ispituju svojstva podataka koja bi potvrdila adekvatnost korištene metode.

Ovaj pak članak pruža odgovore na česta metodološka pitanja u istraživanjima koja se bave predviđanjem stečaja trgovačkih društava, kao što su odabir homogenog slučajnog uzorka, selekcija financijskih i nefinancijskih pokazatelja relevantnih za predviđanje stečaja, odabir optimalne granične vjerojatnosti za klasifikaciju te specifikacija prior distribucija za primjenu Bayesove logističke regresije. Pritom, adekvatno specificirani slabo informativni priori glavni su doprinos istraživanja te posljedično i određivanje razine neizvjesnosti pojedinačnih indikatora stečaja i prediktivne sposobnosti BLR.

2. PREGLED LITERATURE

Znanstvena analiza predviđanja stečaja započela je tijekom 1930-ih godina. Prvo pionirsko istraživanje Fitzpatricka (1932), bilo je usmjereno na analizu pojedinačnih financijskih omjera. On je predložio da bi manje važnosti (težine ili pondere) trebalo pridružiti koeficijentu trenutne likvidnosti i koeficijentu ubrzane likvidnosti ako je poduzeće opterećeno dugoročnim obvezama. Godine 1968. Altman je objavio prve rezultate multivarijantne analize stečaja, što se smatra početkom suvremenog pristupa predviđanju bankrota. Altman je unaprijedio tadašnja istraživanja primjenom diskriminacijske analize, a njegov Z-score model postao je standardna metoda predviđanja bankrota najmanje 20 godina nakon njegova nastanka. Istovremeno

je analizirao pet financijskih omjera na uzorku od 33 zdrava poduzeća i 33 stečajna, pokazujući da se bankrot može predvidjeti za godinu dana unaprijed s točnošću od 95% (Altman, 1968). Beaver (1966) pak zaključuje da je najbolji diskriminator između uspješnih i neuspješnih poduzeća omjer novčanog toka i ukupnog duga, s prediktivnom sposobnošću s rasponom od 78% do 87%, u periodu od 5 do godinu dana prije bankrota. Edmister (1972) koristi *stepwise* multivarijantnu diskriminacijsku analizu s 19 pokazatelja na uzorku od 42 mala poduzeća, razvijajući funkciju sa sedam varijabli koje predviđaju stečaj s točnošću od 93%. Slično istraživanje proveo je i Deakin (1972) obuhvaćajući 64 poslovna subjekta u razdoblju od 1964. do 1970., te koristeći 14 pokazatelja odabranih prema Beaveru. Ohlson (1980) kritizira prethodna istraživanja o predviđanju poslovnog neuspjeha zbog neispunjenih pretpostavki diskriminacijske analize i zavisnih uzoraka. On je analizirao uzorak od 105 bankrotiranih poduzeća i 2058 slučajno odabranih zdravih poduzeća, identificirajući četiri statistički značajna financijska pokazatelja od ukupno devet: veličina, omjer ukupnih obveza i ukupne imovine, omjer zarada nakon kamata i poreza te ukupne imovine i omjer radnog kapitala i ukupne imovine.

Nakon 1980-ih, mnoga istraživanja su napustila upotrebu diskriminacijske analize i sve više se okreću probabilističkim modelima poput probit i logit modela, zbog njihove robusnosti i prednosti (Zmijewski, 1984; Zavrgen, 1985). Modeli se razlikuju i po varijablama koje se koriste za predviđanje te po pristupima modeliranju. Bellovary i sur. (2007) navode najčešće pristupe modeliranju, uključujući diskriminantnu analizu, logit i probit modele, neuronske mreže, evolucijske algoritme i algoritme strojnog učenja. U radu Zenzerovića i Peruška (2021) prikazane su najčešće korištene statističke i matematičke metode za predviđanje stečaja. Fijorek i Grotowski (2012) u istraživanju predikcije bankrota poljskih poduzeća ističu ključne varijable za predviđanje, poput koeficijenta trenutne likvidnosti, operativne marže, neto profitne marže, odnosa kratkoročnih obveza i ukupne imovine te odnosa ukupne imovine i vlasničke glavnice. Primijenili su penaliziranu metodu najveće vjerodostojnosti za procjenu parametara modela logističke regresije (LR). Prema njihovim rezul-

tatima, moguće je predvidjeti bankrot dvije godine unaprijed, s točnošću između 70% i 96,6%. Sfakianakis (2021) je specificirao model s novim skupom prediktorskih varijabli, postižući značajnu diskriminaciju među grčkim proizvodnim tvrtkama. Korišten je uravnotežen uzorak od 28 proizvodnih poduzeća u razdoblju od 2008. do 2015. Koeficijent ubrzane likvidnosti, omjer zaliha i tekućih obveza, pokriće novčanim tokom i ekonomska dodana vrijednost podijeljena s ukupnom imovinom, pokazali su se značajnima za predviđanje bankrota. Korištenjem tih varijabli ispravno je klasificirano 96,43% poduzeća godinu dana prije stečaja. Prilagođeni model predviđanja za dvije i tri godine prije stečaja koristio je iste varijable i ispravno klasificirao 92,86% i 89,29% poduzeća. Također, brojni istraživači diljem svijeta pokušali su predvidjeti bankrot, uključujući Ray (2011), Bhunia i suradnike (2011), Rashid i Abbas (2011), a recentni se pristupi oslanjaju na tehnike strojnog učenja, posebice neuronskih mreža (Sabek, 2023).

Bogdan (2021) je istraživao stečaj u hrvatskoj restoranskoj industriji koristeći dvadeset eksplanatornih varijabli. Zbog problema multikolinearnosti, devet prediktora je eliminirano, a nakon *stepwise* procedure odabira, razvio je model s dvije varijable; koeficijent vlastitog financiranja i EBITDA. U drugom radu, Bogdan i sur. (2021) predviđaju stečaj na temelju ukupne populacije hrvatskih poduzeća, uzorkujući 1099 poduzeća i koristeći 21 financijski pokazatelj. Točnost modela diskriminacije analize (DA) bila je 73,7%, a logit modela 76,3%. Prethodno su Bogdan i sur. (2019) testirali i kalibrirali Altmanov Z-score model na uzorku od 52 poduzeća, postižući uspješnost predviđanja od 73,08% do 84,62% za različite vremenske horizonte predviđanja. Pervan i sur. (2011) istraživali su korisnost javno dostupnih informacija i financijskih izvještaja za razvoj modela predviđanja stečaja. DA model je imao točnost od 79,5% na temelju tekuće likvidnosti, stupnja zaduženosti i EBIT omjera, dok je Pervan (2017) logističkom regresijom utvrdio statistički značajne pokazatelje za predviđanje stečaja malih i srednjih proizvodnih poduzeća. Vareško i Zubin (2022) analizirali su utjecaj pandemije COVID-19 na sektor ugostiteljstva u Istarskoj županiji koristeći prilagođeni Altmanov Z-score model. Zaključili su da je broj tvrtki s visokim rizikom stečaja porastao,

dok se prosječna vrijednost z-scora smanjila zbog pandemije. Šarlija i Jeger (2011) razvili su tri modela logističke regresije kako bi pratili relativnu važnost financijskih pokazatelja tijekom recesije. Pokazatelji aktivnosti su se pokazali značajnima u predviđanju stečaja. Vuković i sur. (2020) procijenili su logit model za predviđanje stečaja poduzeća u veleprodajnom sektoru u zapadnoj Europi, koristeći 23 financijska pokazatelja u razdoblju od 2010. do 2018. Najvažniji prediktori stečaja bili su povrat na kapital, kratkotrajna imovina/ukupna imovina, solventnost, obrtaj radnog kapitala, i zalihe/kratkotrajna imovina.

3. METODOLOGIJA I PODACI

Model standardne logističke regresije s dihotomnom zavisnom varijablom je linearni probablistički model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i$$

$$i = 1, 2, \dots, n \quad (1)$$

$$y_i = \begin{cases} 1 & \text{ako je događaj nastupio (stečaj)} \\ 0 & \text{ako događaj nije nastupio (inače),} \end{cases}$$

Model (1) nije adekvatan jer se metodom najmanjih kvadrata OLS (engl. *Ordinary Least Squares*) očekivane vrijednosti ne mogu ograničiti u zatvorenom intervalu $[0, 1]$, što je nužno budući da je vjerojatnost nenegativna i normirana, tj. $0 \leq p \leq 1$. Također, greške relacije ε_i neće zadovoljiti pretpostavku normalnosti kada zavisna varijabla ima samo dva moguća ishoda. Prikladnije je pretpostaviti da je zavisna slučajna varijabla Bernulijeva, što naposljetku rezultira Binomnom distribucijom vjerojatnosti kada se promatra n nezavisnih opažanja. Shodno tome nelinearna je specifikacija probablističkog modela adekvatnija te podrazumijeva inverznu logit transformaciju jednadžbe (1):

$$P(y_i = 1) = \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i}} \quad (2)$$

Vjerojatnost realizacije određenog događaja $P(y_i = 1)$ je nelinearna monotonno neopadajuća funkcija slijedećeg oblika:

$$p_i = \text{logit}^{-1}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i), \quad (3)$$

pri čemu je e baza prirodnog logaritma, $\beta_0, \beta_1, \dots, \beta_k$ su parametri u modelu koji se procjenjuju na osnovu uzorka n , dok je p_i vjerojatnost stečaja, tj. $p_i = P(y_i = 1)$. Jednadžba (3) predložuje model logističke regresije (LR) u kojem je zavisna varijabla funkcija distribucije vjerojatnosti skupa nezavisnih varijabli logističkog oblika (oblika slova S). Radi jednostavnosti model logističke regresije zapisuje se u obliku:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i, \quad (4)$$

pri čemu je $\log\left(\frac{p_i}{1-p_i}\right)$ logit transformacija, tj. logaritamska vrijednost omjera šansi (engl. *log odds ratio*). Umjesto OLS metode parametri logističke regresije (4) procjenjuju se metodom najveće vjerodostojnosti MLE (engl. *Maximum Likelihood Estimation*). MLE zahtjeva formiranje funkcije vjerodostojnosti koja izražava vjerojatnost promatranih podataka kao funkciju nepoznatih parametara. Doprinos funkciji vjerodostojnosti za svaku vrijednost zavisne varijable y_i i vrijednost prediktora x_i definiran je Bernulijevom distribucijom:

$$P(y = 1) = p_i^{y_i} (1 - p_i)^{1-y_i} \quad (5)$$

$$= \left(\frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right)^{y_i} \cdot \left(1 - \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right)^{1-y_i}.$$

Pretpostavljajući da su opažanja neovisna, funkcija vjerodostojnosti samo je umnožak pojedinačnih doprinosa:

$$L = \prod_{i=1}^n \left(\frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right)^{y_i} \cdot \left(1 - \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right)^{1-y_i} \quad (6)$$

Jednostavnija verzija ove funkcije dobiva se logaritmiranjem funkcije vjerodostojnosti, koja se naziva funkcija log-vjerodostojnosti (engl. *log-likelihood*):

$$\begin{aligned} \text{Log}L = \sum_{i=1}^n y_i \log\left(\frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}\right) \\ \cdot (1 - y_i) \log\left(1 - \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}\right) \end{aligned} \quad (7)$$

Funkcija (7) derivira se s obzirom na β_0 i β_1 te izjednači dva rezultirajuća izraza s nulom kako bi pronašli vrijednost parametara koji maksi-

miziraju gornju funkciju. Koristi se numerička iterativna metoda za rješavanje jednadžbi jer ne postoji analitičko rješenje, a dobivene vrijednosti β_0 i β_1 nazivaju se procjenama najveće vjerodostojnosti istoimenih parametara.

Bayesov pristup pak podrazumijeva upotrebu modela potpune vjerojatnosti kojim se opisuje ne samo neizvjesnost varijable y , uvjetovane skupom nepoznatih parametara β , već i nesigurnost u istim tim parametrima koja se iskazuje njihovom prior distribucijom. Dakle, odnos između priorne vjerojatnosti parametara (prije nego što podaci postanu opaženi) i posteriornih vjerojatnosti istih parametara zapisuje se pomoću Bayesove formule:

$$p(\beta|y) = \frac{p(y|\beta) \cdot p(\beta)}{p(y)}. \quad (8)$$

U izrazu (8) posteriorna vjerojatnost skupa parametara β , uvjetovana opaženim podacima $p(\beta|y)$, jednaka je umnošku funkcije vjerodostojnosti $p(y|\beta)$ i priorne vjerojatnosti $p(\beta)$, podijeljena s marginalnom vjerojatnosti opaženih podataka $p(y)$. Marginalna (granična) vjerojatnost u nazivniku ne daje nikakve informacije o tome koje vrijednosti parametara su više ili manje vjerojatne, stoga se ona zanemaruje u postupku ažuriranja. Skup parametara β , koji se procjenjuje, tretira se kao slučajna varijabla te ako prior distribuciju vjerojatnosti označimo s $f(\beta)$, pretpostavljajući, primjerice, normalnu distribuciju sa sredinom nula 0 i dovoljno velikom varijancom σ_0^2 , tada je

$$f(\beta) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left[-\frac{\beta^2}{2\sigma_0^2}\right], \quad -\infty < \beta < \infty, \quad (9)$$

Pretpostavljajući da su parametri β_j neovisni te da imaju jednake priore, tada je zajednička prior distribucija za sve regresijske koeficijente jednaka umnošku pojedinačnih prior distribucija:

$$f(\beta_0, \beta_1, \dots, \beta_k) = \prod_{j=1}^{k+1} \frac{1}{\sigma_{j0} \sqrt{2\pi}} \exp\left[-\frac{\beta_j^2}{2\sigma_{j0}^2}\right] \quad (10)$$

Potom se formira posteriorna distribucija kao uvjetna funkcija gustoće vjerojatnosti β , temeljeno na Bayesovom teoremu:

$$f(\beta|y_i, x_i) = f(y_i, x_i|\beta) \cdot f(\beta) \quad (11)$$

Uzimajući u obzir jednakost (6) i (9) te funkciju (11) posteriorna distribucija je rezultat množenja prior distribucije preko svih parametara s funkcijom vjerodostojnosti:

$$f(\beta_0, \beta_1, \dots, \beta_k | y_i, x_i, \sigma_{j0}) = \prod_{i=1}^n \left(\frac{e^{e^{\beta_0 + \beta_1 x_i}}}{1 + e^{e^{\beta_0 + \beta_1 x_i}}} \right)^{y_i} \left(1 - \frac{e^{e^{\beta_0 + \beta_1 x_i}}}{1 + e^{e^{\beta_0 + \beta_1 x_i}}} \right)^{1-y_i} \cdot \prod_{j=1}^{k+1} \frac{1}{\sigma_{j0} \sqrt{2\pi}} \exp \left[-\frac{\beta_j^2}{2\sigma_{j0}^2} \right] \quad (12)$$

U jednadžbama (9), (10) i (12) normalna distribucija sa sredinom nula i varijancom σ_0^2 je specificirana kao prior distribucija. Međutim, prior distribucija parametara nužno ne mora biti normalnog oblika niti svaki parametar mora imati istu prior distribuciju. Nadalje, parametri se također procjenjuju maksimizirajući funkciju vjerodostojnosti, ali se ona provodi u velikom broj ponavljanja (engl. *repetitions*) kroz određen broj Markovljevih lanaca (engl. *Markov chain*) pri čemu se koriste Monte Carlo simulacije. Takav je postupak objedinjen pod nazivom Monte Carlo Markovljevi lanci (engl. *Markov Chain Monte Carlo* - MCMC). MCMC je računalno vođena metoda kojom se procjenjuje posteriorna distribucija parametara (Gamerman i Lopes, 2006). Postoji više MCMC algoritma koji se koriste za izvođenje posteriorne distribucije parametara u Bayesovom modelu, a najčešće je u primjeni Metropolis algoritam (Bürkner, 2017).

Prior distribucija može se odrediti na osnovu prethodnih eksperimenata, provedenih istraživanja ili subjektivne procjene istraživača. Općenito, postoje tri vrste priora: (a) informativni prior, koji daje potpune informacije o parametrima, (b) slabo informativni prior, koji ne daje potpune informacije, ali su dovoljno informativne za procjenu parametara i (c) neinformativni prior, koji pridružuje jednaku vjerojatnost svim mogućim vrijednostima parametara, što je specijalni slučaj standardne logističke regresije. U empirijskim istraživanjima slabo informativni prior je najzastupljeniji i to najčešće Studentov *t*-prior jer omogućuje dobro prilagođavanje distribucije vjerojatnosti očekivanim vrijednostima parametara koje mogu poprimiti i ekstremne vrijednosti zbog pojave „debelih repova“

(engl. *fat-tails*) i asimetrije (Gelman i sur. 2008). Upravo se to svojstvo može kontrolirati pomoću stupnjeve slobode iz kojih proizlazi i sama varijanca *t*-distribucije. Preporuka je koristiti *t*-distribuciju, centriranom oko nule, s malim brojem stupnjeva slobode (engl. *degrees of freedom*), iako se za različite stupnjeve slobode mogu dobiti specijalni slučajevi. Primjerice, normalna je distribucija granični slučaj *t*-distribucije s beskonačnim stupnjevima slobode $df=\infty$ koja je simetrična i nema debelih repova ili pak Cauchyjeva distribucija kao granični slučaj *t*-distribucije s jednim stupnjem slobode $df=1$ koja je ekstremno desnostrano asimetrična i ima zadebljani samo desni rep.

Prije deskriptivne statistika odabranih varijabli za previđanje stečaja, u nastavku, je objašnjen izbor slučajnog uzorka poduzeća. Naime, ciljanu populaciju čine sva trgovačka društva u djelatnosti trgovine na veliko i malo (djelatnost G prema NKD-u, osim trgovine motornim vozilima i motociklima) koja su u trogodišnjem razdoblju predali godišnja financijska izvješća (za 2021., 2020. i 2019. godinu) te kojima je poznat status zaključno s 2022. godinom (u stečaju ili nije u stečaju). Prema podacima FINA-e koji su izvezeni pomoću infoBIZ servisa takvih trgovačkih društava ima 17180. Jednostavno slučajno uzorkovanje ne bi bilo reprezentativno jer bi se zbog velike populacije ($N=17180$) mogao dobiti uzorak, od primjerice 2000 trgovačkih društava među kojima ne bi uopće bilo stečajnih društava ili bi ih bilo jako malo pa bi izvedeni zaključci istraživanja bili pristrani ili samo prividni. S obzirom na izrazitu heterogenost ciljane populacije prema veličini trgovačkih društava (mikro ili mali), vrsti poddjelatnosti (trgovina na veliko - poddjelatnost G46 ili trgovina na malo - poddjelatnost G47) i statusu (u stečaju ili nije u stečaju), ista je podijeljena na 8 stratumata (Moćan, 2023). Budući da su po dvije kategorije u svakom od tri navedena obilježja (veličina, poddjelatnost i status), njihovom se kombinacijom došlo se do osam stratumata. Potrebno je naglasiti da ciljanom populacijom ($N=17180$) nisu obuhvaćena srednja i velika trgovačka društva iz dva razloga: a) samo je jedno društvo u stečaju srednje veličine u trgovini na veliko (CELOX d.o.o.) i jedno veliko stečajno društvo u trgovini na malo (NAMA d.d.) stoga nije razumno formirati stratumate s jednom jedinicom promatranja

Tablica 1: Stratificirani uzorak trgovačkih društava u 2022. godini (apsolutni broj # i relativni broj %) prema veličini, poddjelatnosti i statusu

Veličina		Mala trgovačka društva		Mikro trgovačka društva		Ukupno
Poddjelatnost Stečaj		Trgovina na malo	Trgovina na veliko	Trgovina na malo	Trgovina na veliko	
0	#	138	452	693	2051	3334
	%	99,1	99	95,5	96,1	96,5
1	#	1	4	31	66	102
	%	0,7	0,9	4,3	3,1	3,0
Ukupno	#	139	457	726	2134	3456
	%	100	100	100	100	100

Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

i b) uključivanje srednjih i velikih trgovačkih društava koja nisu u stečaju uzrokovalo bi pristranost u rezultatima previđanja stečaja, ali i zbog utjecaja ekstremnih vrijednosti financijskih pokazatelja koje ta društva generiraju (Moćan, 2023). Mikro i mala trgovačka društva dominiraju populacijom s udjelom od 98% pa je izoliranje srednjih i velikih trgovačkih društva iz ciljne populacije bilo logično iz već navedenih razloga. Iz ciljne je populacije izabrano 20% poduzeća proporcionalnim stratificiranim uzorkovanjem, tako da se u uzorku nađe najmanje 100 stečajnih trgovačkih društava u apsolutnom broju što je minimalni zahtjev kada se ista uspoređuju s ne stečajnim trgovačkim društvima kojih u ciljanoj populaciji ima gotovo 33 puta više, tj. 16670 naspram 510 (Moćan, 2023). Struktura stratificiranog uzorka identična je strukturi populacije, a konačan uzorak veličine $n=3456$ sadrži 102 stečajna trgovačka društva (Tablica 1).

Temeljem dostupnih podataka FINA-e i odgovarajućih AOP pozicija financijskih izvještaja izračunati su financijski pokazatelji iz četiri kategorije: likvidnost, zaduženost, aktivnost i profitabilnost. Ovo je istraživanje ograničeno na 10 varijabli koje potencijalno dobro previđaju stečaj trgovačkih društava; 8 financijskih pokazatelja iz navedene 4 kategorije i 2 nefinancijska poka-

zatelja. Više od 10 varijabli se nije razmatralo iz sljedećih razloga: a) financijski pokazatelji, ne nužno iz iste kategorije, mogu biti visoko korelirani te utjecati na pristranost parametara ili čak mogu imati netočan predznak kao posljedica ozbiljnog problema multikolinearnosti i b) ne poštujući načelo parsimonije može se dobiti prividno adekvatan model samo zato što sadrži puno prediktora, a istovremeno se zbog njihove brojnosti može izgubiti veliki broj stupnjeva slobode koji izravno utječu na empirijsku razinu značajnosti (*p-vrijednost*) prilikom testiranja signifikantnosti procijenjenih parametara (Moćan, 2023). Najčešći financijski pokazatelji koji su se u prethodnim istraživanjima pokazali signifikantnim za predviđanje bankrota je koeficijent tekuće likvidnosti, koeficijent zaduženosti i rentabilnost ukupne imovine, stoga su prioritet analize i u ovom istraživanju. Ostali financijski pokazatelji u nešto manjem broju su se pokazali signifikantnima, ali uključeni su u analizu jer kvalitetno odražavaju financijsko stanje sektora trgovine poput trenutne likvidnosti, koeficijenta obrtaja kratkotrajne imovine i ukupne imovine, te EBIT marža i rentabilnost vlastitog kapitala. Tablica 2 predočuje deskriptivno statističke pokazatelje spomenutih varijabli za 2021. godinu za stečajna i ne stečajna trgovačka društva u stratificiranom uzorku.

Tablica 2: Deskriptivna statistika financijskih pokazatelja u 2021. godini

	Stečaj	Prosječna vrijednost	Standardna devijacija	Najmanja vrijednost	Medijalna vrijednost	Najveća vrijednost
Koeficijent trenutne likvidnosti	0	0,32	0,44	0,00	0,12	2,12
	1	0,06	0,21	0,00	0,00	1,84
Koeficijent tekuće likvidnosti	0	1,82	1,11	0,00	1,57	5,10
	1	0,66	0,53	0,00	0,58	2,26
Koeficijent zaduženost	0	0,67	0,34	0,00	0,63	2,06
	1	1,45	0,99	0,00	1,14	4,84
Koeficijent obrtaja ukupne imovine	0	1,51	1,06	0,00	1,41	5,06
	1	0,13	0,30	0,00	0,00	1,57
Koeficijent obrtaja kratkotrajne imovine	0	1,97	1,34	0,00	1,82	6,20
	1	0,18	0,33	0,00	0,00	1,57
Rentabilnost ukupne imovine	0	4,53	7,04	-15,43	3,01	26,10
	1	-2,88	10,56	-46,82	-0,06	26,45
Rentabilnost vlasničke glavnice	0	14,62	15,68	0,00	9,38	67,73
	1	4,04	17,38	0,00	0,00	154,18
EBIT marža	0	3,72	5,03	-10,86	3,02	18,05
	1	-7,16	33,84	-146,23	0,00	90,91

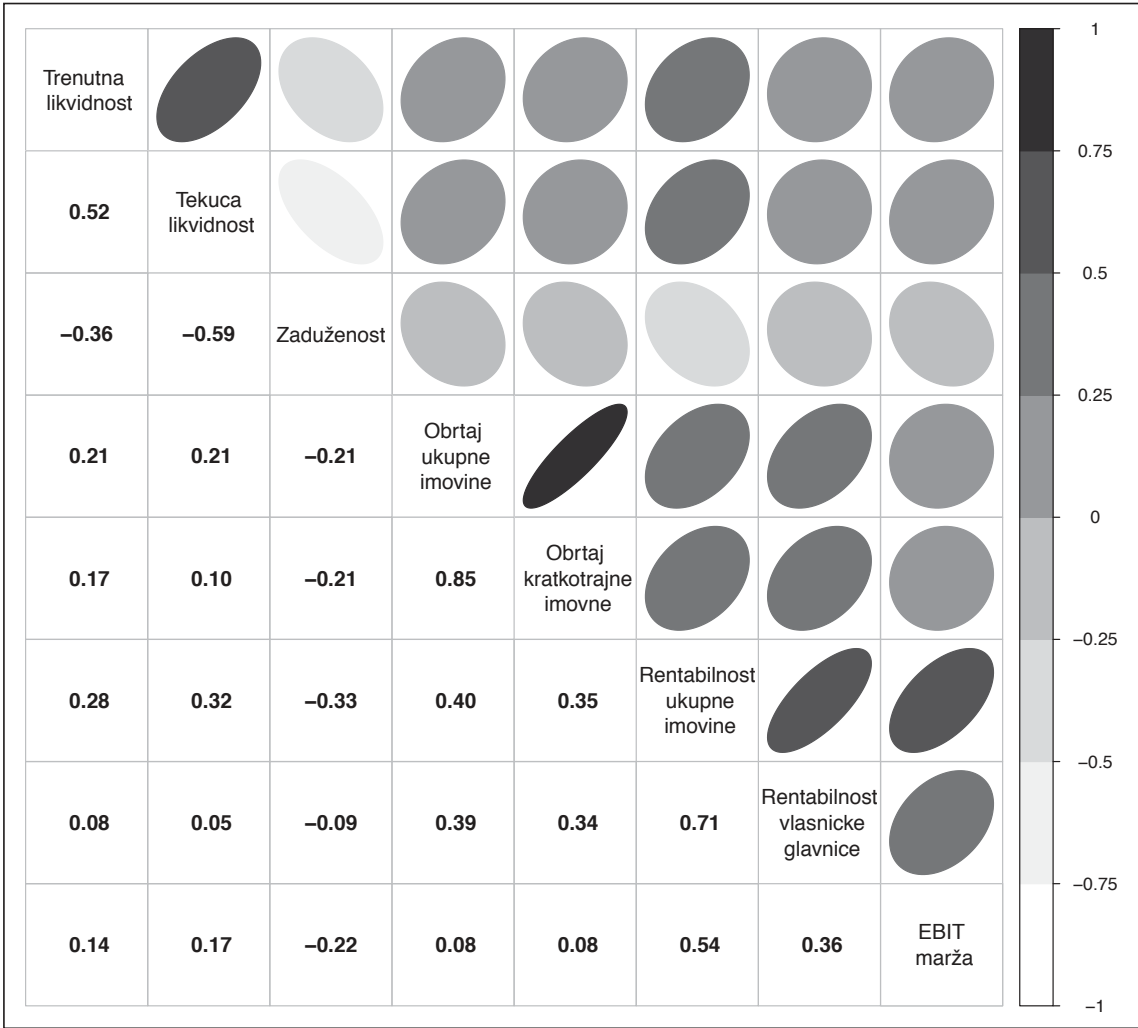
Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

Iz Slike 1. uočava se negativna korelacija jedino između koeficijenta zaduženosti naspram ostalih varijabli, iako umjerenog intenziteta. S vrlo jakim i pozitivnom korelacijom su koeficijent obrtaja ukupne imovine i koeficijent obrtaja kratkotrajne imovine, što je očekivano s obzirom da su oba pokazatelja aktivnosti. Stoga je za pretpostaviti da se u modelu neće naći oba koeficijenta, baš kao ni ROA i ROE budući da je relativno visok pozitivan koeficijent korelacije rentabilnosti ukupne imovine i rentabilnosti vlasničke glavnice, a nešto blaža korelacija od 0,52 postoji između koeficijenta trenutne i tekuće likvidnosti. Zamjećuje se da je koreliranost financijskih pokazatelja više izraženija unutar istih kategorija nego između kategorija pokazatelja.

4. EMPIRIJSKO ISTRAŽIVANJE

Empirijsko istraživanje provedeno je u tri faze. Pomoću Akaikeovog informacijskog kriterija AIC (engl. *Akaike Information Criteria*) izabrani su značajni pokazatelji za predviđanje stečaja pomoću standardne logističke regresije u prvoj fazi. Potom su isti indikatori korišteni u Bayesovoj logističkoj regresiji (BLR) s tri različite specifikacije priora da bi se utvrdilo pridonosi li Bayesov pristup prognostičkoj sposobnosti logističke regresije i koja specifikacija priora je najprikladnija u drugoj fazi, što su ujedno glavni istraživački ciljevi (Moćan, 2023). U posljednjoj, trećoj, fazi uspoređene su jednako specificirane Bayesove logističke regresije koje su procijenje-

Slika 1: Korelacija financijskih pokazatelja u 2021. godini



Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

ne u tri uzastopne godine. Drugim riječima, stečaj trgovačkih društava u 2022. godini predviđa se na temelju vrijednosti istoimenih pokazatelja u 2021., 2020. i 2019. godini. Zanimljivo je da je tim godinama obuhvaćen period COVID-19 pandemije, pa stoga postoje idealni uvjeti za primjenu Bayseove logističke regresije u uvjetima neizvjesnosti (Moćan, 2023). Naposljetku će se posteriornom analizom parametara doći do zaključaka koji su pokazatelji više neizvjesni, a koji manje te kako se mijenjaju razine njihove neizvjesnosti s vremenskim pomakom od 1, 2 i 3 godine.

Inicijalno je procijenjen model standardne logističke regresije (LR) uključujući 8 financijskih pokazatelja, 2 nefinancijska pokazatelja i njihovu interakciju predočenu članom umnoška

(Moćan, 2023). Zajedno s konstantnim članom u prvom modelu LR ukupno je procijenjeno 12 parametara, tzv. puni model. Nefinancijski pokazatelji su veličina i poddjelatnost, za koje su generirane tzv. dummy varijable:

$$veličina_i = \begin{cases} 1 & \text{mikro} \\ 0 & \text{malo} \end{cases},$$

$$poddjelatnost_i = \begin{cases} 1 & \text{trgovina na veliko} \\ 0 & \text{trgovina na malo} \end{cases}$$

Navedene dummy varijable također se pojavljuju u umnošku (*veličina x djelatnost*) da bi se ispitalo postoji li možda utjecaj njihove interakcije na vjerojatnost stečaja. Sve procijenjene vrijednosti parametara izražene su u logaritamskim omjerima šansi (Tablica 3).

Tablica 3: Rezultati cjelokupnog i reduciranog modela standardne logističke regresije

	Cjelokupni model	Reducirani model
Konstantni član	-2,4033 (1,4871)	-1,8292*** (0,3627)
Veličina	0,6581 (1,4583)	
Poddjelatnost	0,4971 (1,5860)	
Veličina × djelatnost	-0,6358 (1,6054)	
Trenutna likvidnost	-0,7597 (0,6409)	
Tekuća likvidnost	-0,4715* (0,1947)	-0,5439** (0,1866)
Zaduženost	1,1957*** (0,2348)	1,2267*** (0,2350)
Obrtaj ukupne imovine	-0,1557 (0,9811)	
Obrtaj kratkotrajne imovine	-2,3260** (0,8283)	-2,5097*** (0,3278)
Rentabilnost ukupne imovine	-0,0242 (0,0213)	
Rentabilnost vlasničke glavnice	0,0287** (0,0099)	0,0267** (0,0095)
EBIT marža	-0,0291*** (0,0087)	-0,0335*** (0,0080)
Num.Obs.	3456	3456
Log.Lik.	-260,420	-261,998
AIC	544,8	536,0
RMSE	0,18	0,18
F	14,734	32,622

Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

Standardne greške u zagradama

Razine značajnosti: * $p < 0,10$, ** $p < 0,05$ i *** $p < 0,01$

U reduciranom modelu preostale su varijable koje su izabrane pomoću AIC kriterija i koje su sve statistički značajne, a očigledno je reducirana-

ni model prikladniji naspram punog modela jer ima manju vrijednost AIC. Akaikeov informacijski kriterij je osmišljen kako bi pronašao model koji najviše objašnjava varijacija u podacima, dok istovremeno penalizira modele koji koriste preveliki broj parametara (varijabli) u odnosu na veličinu uzorka (Akaike, 1974). Pomoću navedenog kriterija varijable se iterativno (postupno) dodaju i uklanjaju iz modela logističke regresije sve dok se ne pronađete skup varijabli s najnižom vrijednosti AIC. Upravo je ovaj kriterij reduciranja varijabli korišten u istraživanju, čime se ujedno ublažava problem multikolinearnosti. Reducirani se model bolje prilagođava podacima, bez obzira što ima jednaku vrijednost prosječne pogreške RMSE (engl. *Root Mean Square Error*) te zbog visoke vrijednosti F-testa i pripadajuće niske p -vrijednosti dodatno se zaključuje da je barem jedna varijabla značajna u modelu (Moćan, 2023). Svi predznaci procijenjenih parametara su očekivani i smisleni osim predznaka uz varijablu ROE kao posljedica visoke koreliranosti s pokazateljem ROA. Kako bi se navedeni problem riješio varijabla ROE se neće dalje analizirati, a preostale varijable su zapravo pojedini pokazatelji iz svake kategorije likvidnosti, aktivnosti, zaduženosti i profitabilnosti. Nefinancijske varijable nisu se pokazale statistički značajnima i one se također neće dalje analizirati.

Bayesova logistička regresija, kojom se predviđa stečaj u 2022. godini, procijenit će se u 2021. godini s tri različito specificiranim priornim distribucijam, a potom će se uz jednako adekvatne specificirane priore usporediti s Bayesovom logističkom regresijom u 2020. i u 2019. godini. Tri su specifikacije slabo informativnog priora korištene u Bayesovoj logističkoj regresiji:

- I. za sve parametre normalni prior sa sredinom 0 i standardnom devijacijom 2,5 na ljestvici logaritamskih omjera šansi,
- II. za konstantni član normalni prior sa sredinom -3,47 i standardnom devijacijom 1,74 te za preostale parametre t-prior sa 7 stupnjeva slobode, sredinom 0 i standardnom devijacijom 2,5 na ljestvici logaritamskih omjera šansi i
- III. za konstantni član t-prior s 1 stupnjem slobode, sredinom -3,47 i standardnom devija-

cijom 1,74 te za koeficijente smjera t-prior sa 7 stupnjeva slobode, sredinom 0 i standardnom devijacijom 2,5 na ljestvici logaritamskih omjera šansi.

Prva specifikacija priora je standardna normalna distribucija za sve parametre sa sredinom 0 i standardnom devijacijom 2,5 na ljestvici logaritamskih omjera šansi, pa se model Bayesove logističke regresije može formalno zapisati ovako:

$$y_i | \beta_0, \beta_1, \dots, \beta_4 \sim B(1, p_i)$$

$$\log \left(\frac{p_i}{1 - p_i} \right) = \beta_0 + \sum_{j=1}^4 \beta_j x_{ij}$$

$$\beta_0 \sim N(0; 2,5)$$

$$\beta_j \sim N(0; 2,5)$$
(13)

Ostale postavke u modelu, ključne za konvergenciju Metropolis algoritma, su broj Markovljevih lanaca (4) i broj ponavljanja (5000). Metodom MCMC će se na slučajan način generirati 20000 koeficijenata logističke regresije β_j i to za svako opažanje, pretpostavljajući normalnu distribuciju vjerojatnosti (Moćan 2023). To znači da će generirane vrijednosti β_j biti iz intervala od -2,5 do +2,5 na ljestvici logaritamskih omjer šansi, odnosno od 0 do približno 12 na ljestvici omjera šansi. Antilogaritamska vrijednost koeficijenta β_j od 0 do 1 znači da varijabla x_j smanjuju vjerojatnost stečaja, a vrijednost veća od 1 znači da varijabla x_j povećava vjerojatnost stečaja i to maksimalno do 12 puta što je približno 1100%. Dakle, ako se postavi da je sredina normalnog priora 0 onda se dozvoljavaju negativni i pozitivni utjecaji varijable x_j na vjerojatnost stečaja, dok se standardnom devijacijom na ljestvici logaritamskih omjera šansi kontrolira koliki taj utjecaj može biti, što zapravo izražava naše uvjerenje o tom parametru. Što je veća standardna devijacija prior je slabije informativan, a i razina neizvjesnosti uključena u generiranje posteriorne distribucije parametra β_j je veća, koja naposljetku ovisi ne samo o prioru već i funkciji vjerodostojnosti temeljem opaženih podataka. Isti je prior specificiran i za konstantni član β_0 . Međutim, takva specifikacija priora konstantnog člana nije prikladna. Konstantni član pokazuje vjerojatnost stečaja trgo-

vačkog društva kada bi svi financijski pokazatelji bili nula, tj. razinu vjerojatnosti stečaja bez obzira koje će vrijednosti poprimiti financijski pokazatelji. Iz normalnog priora sa sredinom 0 i standardnom devijacijom 2,5 proizlazi da navedena bezuvjetna vjerojatnost može biti od 0 do približno 92,3% ($\exp(2,5)/(1+\exp(2,5))=12/(1+12)100\%$), što nije suvislo jer je u populaciji, baš kao i u stratificiranom uzorku, stečajnih trgovačkih društava približno 3% (102/3456). Pema tome, distribuciju priora za konstantni član bi trebalo izmijeniti, što je i učinjeno u drugoj specifikaciji (II). Vjerojatnosti od 3% odgovara logaritamska vrijednost omjera šansi od -3,47, a za standardnu devijaciju se uzme polovica tog iznosa pa je to raspon $-3,47 \pm 1,74$. Navedeni raspon odgovara bezuvjetnoj vjerojatnosti stečaja od približno 0% do 15%. Tako specificiran prior je također slabo informativan. U trećoj verziji (III) normalni je prior zamijenjen t-priorom, ali različitim stupnjevima slobode za konstantni član i za regresijske koeficijente (Moćan, 2023). Prednost Studentovog t-priora je što dozvoljava da se u neizvjesnost inkorporiraju i ekstremne vrijednosti koje se na krajevima distribucije odražavaju kao zadebljani repovi. Što su stupnjevi slobode manji t-prior ima zadebljanije repove. Štoviše, za konstantni član specificiran je t-prior s jednim stupnjem slobode kao specijalni slučaj Cauchyjeve distribucije koja osim ekstremnih vrijednosti dozvoljava i asimetričnost, tj. veću koncentriranost distribucije pri niskim vjerojatnostima što je i adekvatno za modeliranje bezuvjetne vjerojatnosti stečaja. Na osnovu generiranih posteriornih distribucija MCMC simulacijom u Tablici 4 prezentirane su prosječne vrijednosti procijenjenih parametara za tri specifikacije Bayesove logističke regresije u 2021. godini (BLR21(I), BLR21(II) i BLR21(III)).

Neke se prosječne vrijednosti procijenjenih parametara pak umjereno razlikuju s obzirom na različite specifikacije prior distribucija. Primarno će se utvrditi koji je model Bayesove logističke regresije u 2021. godini prikladniji s obzirom na tri specifikacije priora, a potom će se za adekvatan model posteriorne distribucije usporediti zajedno na jednom grafikonu u svrhu zaključivanja o nesigurnosti pojedinačnih parametara. Tek će se naposljetku prosječni omjeri šansi, temeljem posteriornih distribucija, interpretirati na tradicionalan način.

Tablica 4: Rezultati Bayesove logističke regresije s različito specificiranim prior distribucijama u 2021. godini

	BLR21(I)	BLR21(II)	BLR21(III)
Konstantni član	-1,802 (0,355)	-1,809 (0,354)	-1,786 (0,351)
Tekuća likvidnost	-0,551 (0,181)	-0,544 (0,173)	-0,552 (0,182)
Zaduženost	1,218 (0,231)	1,217 (0,231)	1,218 (0,226)
Obrtaj kratkotrajne imovine	-2,066 (0,257)	-2,026 (0,251)	-2,113 (0,270)
EBIT marža	-0,029 (0,007)	-0,029 (0,007)	-0,029 (0,007)
Num.Obs.	3456	3456	3456
R ²	0,363	0,360	0,365
Log.Lik.	-265,973	-266,077	-265,960
ELPD	-271,2	-271,3	-271,4
ELPD s.e.	20,6	20,7	20,5
LOOIC	542,5	542,7	542,6
LOOIC s.e.	41,1	40,9	41,3
WAIC	542,4	542,7	542,6
RMSE	0,18	0,18	0,18

Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

Standardne greške u zagradama

U Tablici 4 nema indikatora kojim se testira značajnost prosječnih vrijednosti parametara temeljem *p-vrijednosti*, iako se u zagradama prikazuju standardne pogreške njihovih posteriornih distribucija. Kod Bayesovog pristupa značajnost parametra se ne testira kao kod frekventističkog pristupa, ali se mogu donijeti zaključci koje su varijable više ili manje neizvjesne. Do ovih informacija ne može se doći pomoću standardne logističke regresije. Nadalje, log-vjerojatnost

može se koristiti za izračunavanje LOOIC informacijskog kriterija (engl. *leave one out information criterion*) koji se temelji na konceptu križne validacije (Vehtari i sur. 2017) i asimptotski je jednak Watanabe-Akaike informacijskom kriteriju WAIC (engl. *Watanabe-Akaike information criterion*) te će služiti za usporedbu modela BLR21(I), BLR21(II) i BLR21(III). Niža vrijednost LOOIC i WAIC ukazuju na bolju prikladnost modela (Bürkner, 2017). U Tablici 4 prikazani su LOOIC i WAIC za svaki model s odgovarajućom standardnom pogreškom, te očekivana vrijednost prediktivne gustoće ELPD (engl. *Expected Log Predictive Density*).

Uz jednaku prognostičku pogrešku od 0,18 u svim modelima, modelom BLR21(III) je protumačeno 36,5% varijacija stečaja što je za 0,2% više od BLR21(I) i 0,5% više od BLR21(II). Prema LOOIC i WAIC kriterijima sva tri modela su gotovo jednako prikladna, dok je ELPD modela BLR21(III) samo za 0,1 veći od BLR21(II) te za 0,2 veći od modela BLR21(I), što ukazuje da ne postoji značajna superiornost posteriorne prediktivne sposobnosti jednog modela naspram ostalih modela u kontekstu interne valjanosti. Ipak, BLR21(III) se bolje prilagođava podacima s obzirom na log-vjerodostojnost te se zaključuje da taj model ima veću prediktivnu sposobnost s većim ELPD-om (Moćan, 2023). Prior distribucija za konstantni član modela BLR21(III) specificirana je kao Studentova t-distribucija s 1 stupnjem slobode (specijalni slučaj Cauchyjeve distribucije) te sredinom -3,47 i standardnom devijacijom 1,74, dok je za ostale koeficijente specificirana Studentova t-distribucija sa 7 stupnjeva slobode, sredinom 0 i standardnom devijacijom 2,5 na ljestvici logaritamskih omjera šansi. Taj se model zapisuje:

$$y_i | \beta_0, \beta_1, \dots, \beta_4 \sim B(1, p_i)$$

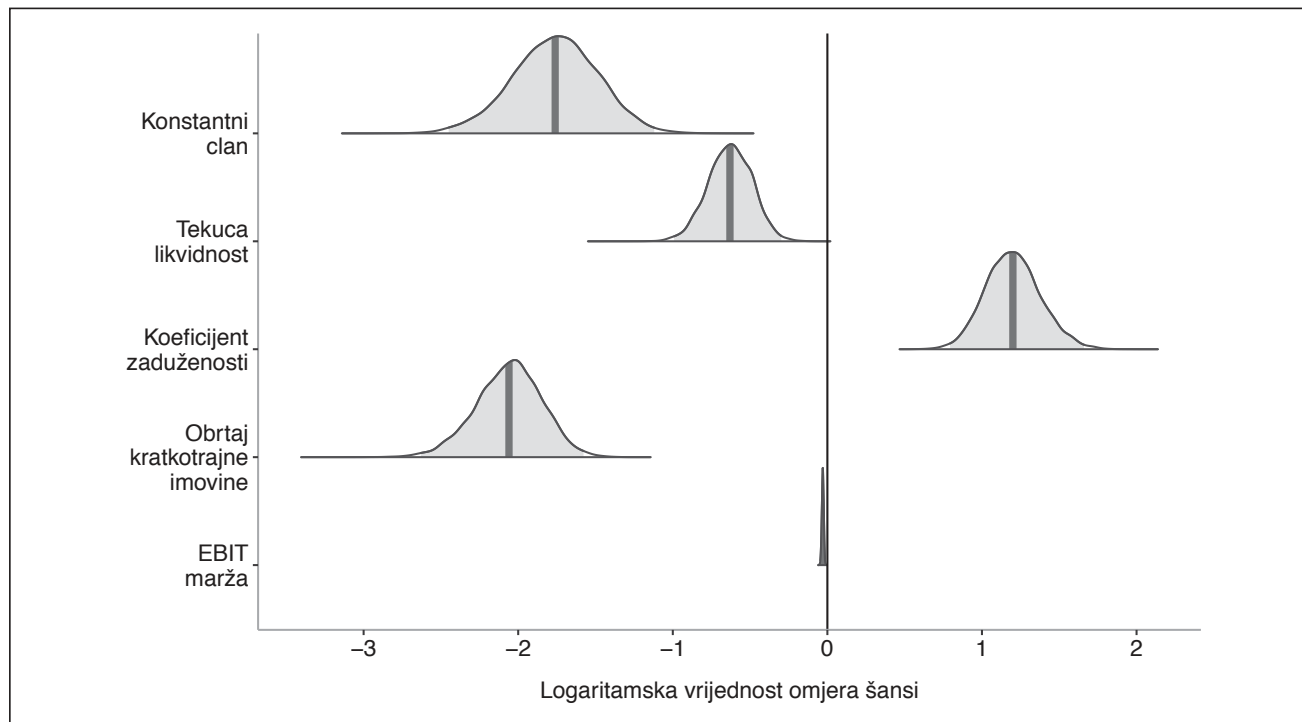
$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \sum_{j=1}^4 \beta_j x_{ij} \quad (14)$$

$$\beta_0 \sim t_{df=1}(-3,47; 1,74)$$

$$\beta_j \sim t_{df=7}(0; 2,5)$$

Rezultati MCMC simulacije su posteriorne distribucije parametara uz svaku nezavisnu vari-

Slika2: Posterior distribucije parametara Bayesove logističke regresije za predviđanje stečaja (financijski pokazatelji iz 2021. godine)



Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

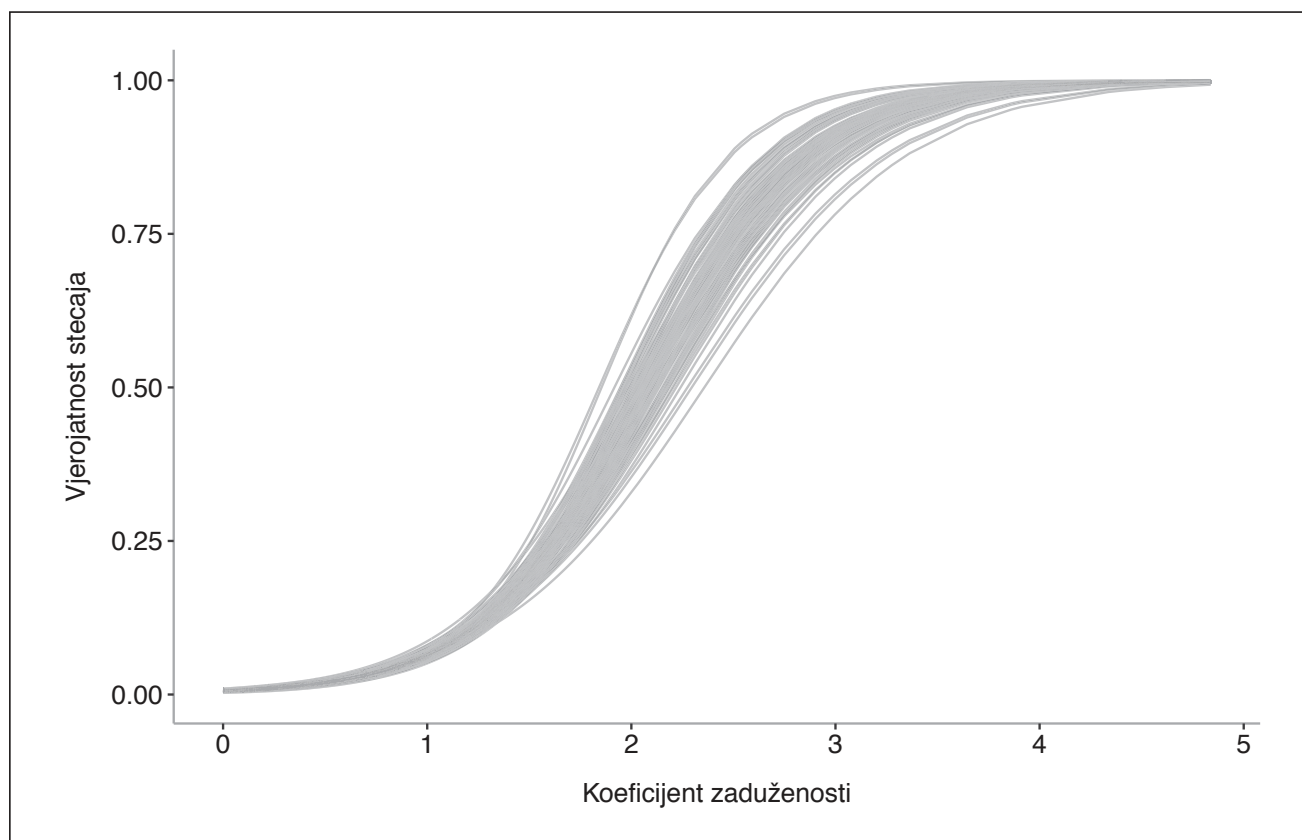
jablu pojedinačno, koje su uspoređene na istom grafikonu (Slika 2).

Posteriorne distribucije tekuće likvidnosti, obrtaja kratkoročne imovine, EBIT marže i konstantnog člana pozicionirane su lijevo od nule, što znači da smanjuju vjerojatnost stečaja jer su prosječne logaritamske vrijednosti omjera šansi negativne. Jedino je posteriorna distribucija logaritamskih omjera šansi zaduženosti položena desno od nule što znači da jedino ona doprinosi stečaju poduzeća (Moćan, 2023). Također, što je posteriorna distribucija udaljenija od ishodišta to promatrana varijabla ima veću diskriminatornu snagu, tj. bolje objašnjava razlike između poduzeća u stečaju od ne stečajnih poduzeća. To je u ovom slučaju primjene koeficijent obrtaja kratkoročne imovine. Iako EBIT marža najmanje diskriminira stečajna od ne stečajnih poduzeća ona je naspram ostalih varijabli „najsigurnija“, tj. najmanje neizvjesna budući da pripadajuća posteriorna distribucija ima najmanju varijancu iskazanu u logaritamskim omjerima šansi, a distribucija je uža i više koncentrirana oko prosječne vrijednosti. Najneizvjesnija je varijabla obrtaja kratkoročne imovine s umjereno

izraženom ljevostranom asimetrijom, dok su zaduženost i tekuća likvidnost umjereno i jednako neizvjesne varijable.

Temeljem modela BLR21(III) može se zaključiti da su šanse za stečaj manje za 19,5% do 60,3% ako se koeficijent tekuće likvidnosti u prosjeku poveća za 1 uz nepromijenjenost ostalih financijskih pokazatelja (95%-tni posteriorni prediktivni interval). Šanse za stečaj smanjuju se za 80,4% do 93,3% s jediničnim povećanjem koeficijenta obrtaja kratkotrajne imovine te između 1,5% i 4,3% s povećanjem EBIT marže za 1%. Nasuprot tome, povećanje koeficijenta zaduženosti za 1 povećat će u prosjeku šanse stečaja između 121,1% i 439,7%.

Slika 3 prikazuje 100 generiranih logističkih krivulja od njih ukupno 20000. One ilustriraju nesigurnost oko povećanja vjerojatnosti stečaja s porastom zaduženosti. Primjerice, ako je koeficijent zaduženosti 2, vjerojatnost stečaja je najvjerojatnije između 37,5% i 62,5% tj. ima toliki posteriorni interval. Veći posterirni interval znači veću neizvjesnost.

Slika 3: Simulacija odnosa vjerojatnosti stečaja i koeficijenta zaduženosti za 2021. godinu

Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

U Tablici 5 uspoređene su procijenjene prosječne vrijednosti parametara izražene u logaritamskim omjerima šansi modela Bayesove logističke regresije s jednako specificiranim prior distribucijama u 2021. 2020. i 2019. godini.

EBIT marža najmanje diskriminira stečajna poduzeća od aktivnih u 2021. godini, no s odmakom vremena, u 2020. i 2019. godini u potpunosti gubi diskriminatornu snagu te nema utjecaj na vjerojatnost stečaja. Suprotno se događa s koeficijentom tekuće likvidnosti, njegova diskriminatorna snaga umjereno raste i najveća je u 2019. godini. Neizvjesnost varijable tekuće likvidnosti prvo u 2020. godini pada, a potom raste u 2019. godini. Koeficijent obrtaja kratkoročne imovine i koeficijent zaduženosti s odmakom vremena u potpunosti drugačije diskriminiraju poduzeća. Koeficijent obrtaja kratkoročne imovine ima najveću diskriminatornu snagu u 2021. godini, ali je i najneizvjesnija varijabla zato što pripadajuća posteriorna distribucija ima najmanju varijancu iskazanu u logaritamskim omjeri-

ma šanse. Odmakom vremena koeficijent obrtaja kratkoročne imovine smanjuje šanse stečaja, gubi diskriminatornu snagu u 2020. godini, da bi u 2019. godini on povećavao šanse za stečaj. Isto se događa s koeficijentom zaduženosti, samo obratno (Slike 4-5).

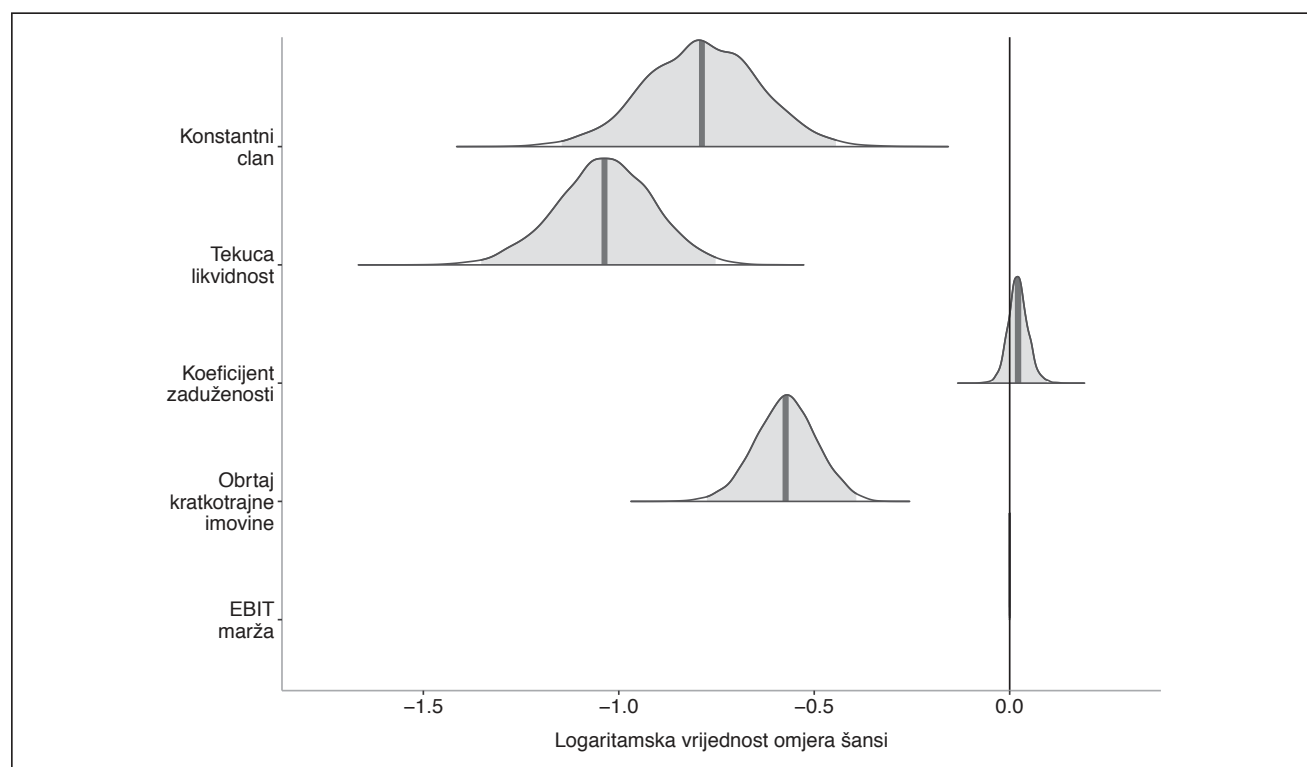
Točnost predviđanja stečaja modelima Bayesove logističke regresije zahtijevaju drugačiji pristup prilikom formiranja klasifikacijske matrice. Naime, zbog upotrebe MCMC metode generirano je 20000 klasifikacijskih matrica. Ako se uzme njihov prosjek postoji velika mogućnost da su rezultati pristrani zbog prekomjernog prilagođavanja podacima (engl. *overfitting*) s obzirom na veliki broj ponavljanja (Vehtari i sur., 2017). Da bi se to izbjeglo klasifikacijska se matrica formira križnom validacijom (engl. *cross-validation*). Križna validacija podrazumijeva evaluaciju prediktivne sposobnosti modela nasumičnim uzorkovanjem klasifikacijskih matrica na one za treniranje (engl. *training*) i na one za testiranje (engl. *testing*) modela. Ovaj postupak se ponavlja

Tablica 5: Rezultati Bayesove logističke regresije s jednako specificiranim prior distribucijama u 2021., 2020. i 2019. godini

	BLR21(III)	BLR20(III)	BLR19(III)
Konstantni član	-1,786 (0,351)	-1,698 (0,263)	-1,647 (0,186)
Tekuća likvidnost	-0,552 (0,182)	-0,716 (0,150)	-1,058 (0,164)
Zaduženost	1,218 (0,226)	0,597 (0,138)	-0,011 (0,011)
Obrtaj kratkotrajne imovine	-2,113 (0,270)	-0,623 (0,100)	0,003 (0,004)
EBIT marža	-0,029 (0,007)	0,000 (0,000)	0,000 (0,000)
Num.Obs.	3456	3456	3456
R2	0,365	0,122	0,042
Log.Lik.	-265,960	-374,455	-422,402
ELPD	-271,4	-393,7	-430,5
ELPD s.e.	20,5	40,0	29,8
LOOIC	542,6	787,3	860,9
LOOIC s.e.	41,3	79,9	59,5
WAIC	542,6	784,2	857,5
RMSE	0,18	0,20	0,21

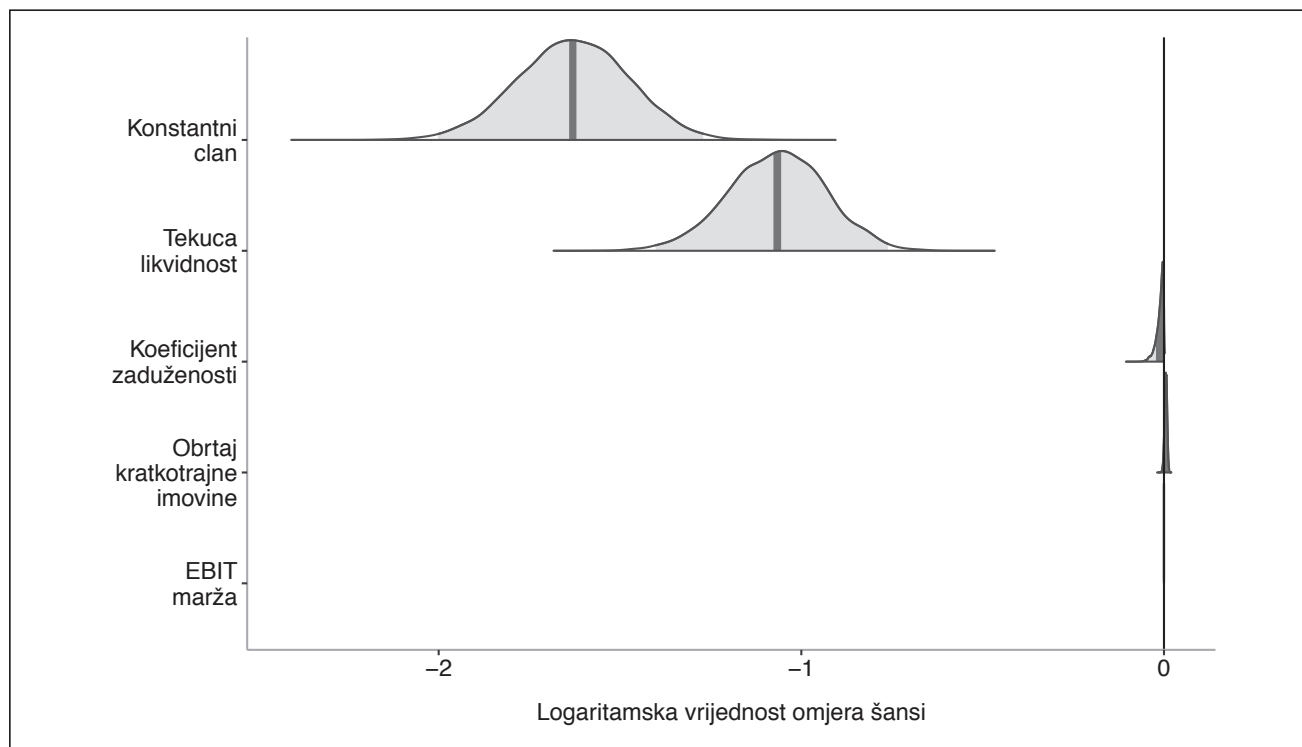
Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e
Standardne pogreške u zagradama

Slika 4: Posterior distribucije parametara Bayesove logističke regresije za predviđanje stečaja (financijski pokazatelji iz 2020. godine)



Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

Slika 5: Posterior distribucije parametara Bayesove logističke regresije za predviđanje stečaja (financijski pokazatelji iz 2019. godine)



Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

Tablica 6: Usporedba prediktivne sposobnosti po godinama predviđanja

Model	Godina predviđanja	Specifičnost	Osjetljivost	Točnost klasifikacije	Optimalna razina praga
Standardna logistička regresija	2021	0,9954	0,21	0,9001	0,125
Bayesova logistička regresija	2021	0,9961	0,2163	0,9022	0,12
Bayesova logistička regresija	2020	0,9896	0,1501	0,8747	0,1
Bayesova logistička regresija	2019	0,9851	0,075	0,7645	0,07

Izvor: izrada autora u programu RStudio prema podacima infoBIZ servisa FINA-e

sve dok svaki od k podskupova ne posluži kao validacijski skup. Križno validirana prediktivna sposobnost, na primjer, prosjek je svih deset točnosti ($k=10$) postignutih na validacijskim skupovima (Tablica 6).

Iz Tablice 6 je evidentno da optimalna razina praga nije jednaka po godinama; ona se sma-

njuje s odmakom vremena. Smanjenjem praga vjerojatnosti, model će više poduzeća točno klasificirati kao stečajna. To vodi većem broju lažno pozitivnih među aktivnim poduzećima, te se time smanjuje specifičnost, a povećava osjetljivost. No u ovom slučaju odmakom vremena i smanjenjem praga vjerojatnosti dolazi do beznačajnog povećanja lažno pozitivnih, a

značajnog povećanja lažno negativnih. Što znači da nepromijenjenim pragom vjerojatnosti bi beznačajno povećala specifičnost, a dodatno bi se narušila osjetljivost. Osjetljivost se pak značajna smanjuje s odmakom od 2 i 3 godine, dok je specifičnost postojana, što je i očekivano jer se u turbulentnim i kriznim razdobljima, koje je i obuhvaćeno ovim istraživanjem, negativni učinci na trgovačka društva događaju jako brzo (s pomakom od jedne godine).

5. ZAKLJUČAK

Točnost predviđanja stečaja uvelike ovisi o metodologiji koja se primjenjuje te o odabiru najprikladnijih prediktora. Premda su metode i modeli za predviđanje stečaja jako raznoliki, primjerice, univarijatna i multivarijatna diskriminacijska analiza, logit i probit modeli, neuronske mreže, ekspertni sustavi, potporni vektori i gradijentna metoda, impresivno su se razvijali posljednjih godina s namjerom poboljšanja njihove interne valjanosti (na uzorku poduzeća za treniranje odnosno procjenu) i eksterne valjanosti (na uzorku poduzeća za validaciju odnosno prognozu). Iako je u ovom radu uzorak poduzeća za procjenu modela jednak uzorku za validaciju, podaci su korišteni s vremenskim pomakom od jedne, dvije i tri godine s ciljem predviđanja stečaja u uvjetima neizvjesnosti. Upravo zato je uspostavljen model standardne logističke regresije (frekventistički pristup) i model Bayesove logističke regresije s tri različite specifikacije priora, koji se naposljetku uspoređuju.

Bayesov pristup logističkoj regresiji pretpostavlja da su parametri slučajni, a varijable fiksne, stoga je on koristan kada postoji velika razina nesigurnosti u podacima. Budući da je sektor trgovine u 2020. godini bio najviše pogođen COVID-19 pandemijom (25,43% svih stečajeva u toj godini je iz sektora trgovine na veliko i malo) glavni cilj rada je bio istražiti pridonosi li Bayesov pristup prediktivnoj sposobnosti logističke regresije u razdoblju neizvjesnosti, tj. u razdoblju COVID krize u kojem je istoimena metoda i primijenjena. Fleksibilnost Bayesove logističke regresije ostvarena je primjenom računalno vođene metode uzorkovanja Markovljevih lanaca Monte Carlo (MCMC), koja se temelji na Bayesovom pravilu za ažuriranje posterior distribucije

parametara s novim podacima i prethodno definiranom prior distribucijom. Pritom je korišteno osam financijskih i dva nefinancijska pokazatelja. Financijski pokazatelji koji su se pokazali značajnima za predviđanje stečaja trgovačkih društava su koeficijent tekuće likvidnosti, koeficijent zaduženosti, koeficijent obrtaja kratkotrajne imovine i EBIT marža. Pritom, koeficijent obrtaja kratkotrajne imovine najbolje razdjeljuje skupinu aktivnih poduzeća od skupine stečajnih poduzeća, iako je najviše neizvjestan. S druge strane, EBIT marža pokazuje manju diskriminacijsku moć, ali je „najizvjesnija“ varijabla, dok su zaduženost i tekuća likvidnost umjereno neizvjesne varijable. Do navedenih se zaključaka došlo izborom prikladnih prior distribucija parametara, koje su sve izvedene na osnovu Studentove t-distribucije, što je značajna doprinos ovog istraživanja. Navedena je distribucija, iako slabo informativna, od velike važnosti za modeliranje neizvjesnosti u parametrima. U tom kontekstu, ključni su stupnjevi slobode kojima se kontrolira razina neizvjesnosti te standardna devijacija na ljestvici logaritamskih omjera šansi. Prikladnost Cauchyjevog priora (specijalnog slučaja t-distribucije s 1 stupnjem slobode) za konstantni član je potvrđena, dok je za preostale koeficijente t-prior sa 7 stupnjeva slobode i standardnom devijacijom od 2,5 na ljestvici logaritamskih omjera šansi, primjenjiviji. Naposljetku, točnost predviđanja stečaja trgovačkih društava u sektoru trgovine na veliko i malo s odmakom od jedne godine je 90,22% Bayesovom logističkom regresijom. Dvije godine prije stečaja trgovačkih društava model je predvidio stečaj s točnošću od 87,47%, dok je tri godine prije stečaja postigao točnost od 76,45%. Veća prediktivna sposobnost Bayesove logističke regresije sugerira njezinu prikladnost u uvjetima neizvjesnosti, naspram standardne logističke regresije.

Literatura

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723. <https://doi.org/10.1109/tac.1974.1100705>
- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of*

- Finance*, 23(4), 589-609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>
- Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, 4, 71-111. <https://doi.org/10.2307/2490171>
- Bellovary, J., Giacomino, A. & Akers, M. (2007). A Review of Bankruptcy Prediction Studies: 1930 to Present. *Journal of Financial Education*, 33, 1-42. <https://www.jstor.org/stable/41948574>
- Bhunja, A., Khan, I. & Mukhuti, S. (2011). Prediction of Financial Distress - A Case Study of Indian Companies. *Asian Journal of Business Management*, 3(3), 210-218. <http://maxwellsci.com/print/ajbm/v3-210-218.pdf>
- Bogdan, S. (2021). Predviđanje stečaja u restoranskoj industriji Hrvatske. *Ekonomski misao i praksa*, 30(1), 99-119. <https://doi.org/10.17818/emip/2021/1.5>
- Bogdan, S., Šikić, L. & Bareša, S. (2021). Predviđanje stečaja na temelju ukupne populacije Hrvatskih poduzeća. *Ekonomski pregled*, 72(5), 643-669. <https://doi.org/10.32910/ep.72.5.1>
- Bogdan, S., Bareša, S. & Hađina, V. (2019). Testiranje primjenjivosti Altmanovog Z-score modela za predviđanje stečaja u Republici Hrvatskoj. *Notitia - Journal for Economic, Business and Social Issues*, 5(1), 31-46. <https://doi.org/10.32676/n.5.1.4>
- Bürkner, P. C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1-28. <https://doi.org/10.18637/jss.v080.i01>
- Deakin, E. B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 10(1), 167-179. <https://doi.org/10.2307/2490225>
- Edmister, R. O. (1972). An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. *The Journal of Financial and Quantitative Analysis*, 7(2), 1477-1493. <https://doi.org/10.2307/2329929>
- Fijorek, K. & Grotowski, M. (2012). Bankruptcy Prediction: Some Results from a Large Sample of Polish Companies. *International Business Research*, 5(9), 70-77. <https://doi.org/10.5539/ibr.v5n9p70>
- Fitzpatrick, P. J. (1932). A Comparison of Ratios of Successful Industrial Enterprises with Those of Failed Companies. *Certified Public Accountant*, 3, 598-605.
- Gamerman, D. & Lopes, H. F. (2006). *Markov chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. New York: Chapman and Hall/CRC. <https://doi.org/10.1201/9781482296426>
- Gelman, A., Jakulin, A., Pittau, M. G. & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2(4), 1360-1383. <https://doi.org/10.1214/08-aos191>
- Keglevi, K. S., Šuker, M. & Novak, E. (2022). Trgovina na malo Republike Hrvatske u vrijeme COVID-19. *CroDiM: International Journal of Marketing Science*, 5(1), 253-262. <https://hrcak.srce.hr/275598>
- Moćan, M. (2023). *Bayesova logistička regresija za predviđanje stečaja trgovačkih društava u djelatnosti trgovine na veliko i malo* (Diplomski rad). Zagreb: Sveučilište u Zagrebu, Ekonomski fakultet. <https://urn.nsk.hr/urn:nbn:hr:148:259183>
- Novak, B. & Sajter, D. (2007). Causes of Bankruptcy in Europe and Croatia. MPRA paper 5833, 1-17. <https://mpra.ub.uni-muenchen.de/5833/>
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109-131. <https://doi.org/10.2307/2490395>
- Ooghe, H. & De Prijcker, S. (2008). Failure processes and causes of company bankruptcy: A typology. *Management Decision*, 46(2), 223-242. <https://doi.org/10.1108/00251740810854131>
- Pervan, I., Pervan, M. & Vukoja, B. (2011). Prediction of company bankruptcy using statistical techniques – case of Croatia. *Croatian Operational Research Review*, 2(1), 158-167. <https://hrcak.srce.hr/96660>
- Pervan, I. (2017). Predviđanje stečaja – SME proizvodna poduzeća u Hrvatskoj. *Zbornik Radova Veleučilišta u Šibeniku*, 10(3-4), 33-45. <https://hrcak.srce.hr/184255>
- Rashid, A. & Abbas, Q. (2011). Predicting Bankruptcy in Pakistan. *Theoretical and Applied Economics*, 18(9), 103-128. <http://www.store.ectap.ro/articole/640.pdf>
- Ray, S. (2011) Prediction of corporate financial distress in glass and glassware sector

- of India. *International Journal of Business Excellence*, 4(6), 678-695. <https://doi.org/10.1504/ijbex.2011.042905>
- Sabek, A. (2023). Unveiling the diverse efficacy of artificial neural networks and logistic regression: A comparative analysis in predicting financial distress. *Croatian Review of Economic, Business and Social Statistics*, 9(1), 16-32. <https://doi.org/10.2478/crebss-2023-0002>
- Sajter, D. (2014). Stečaj: okvir za malverzacije ili za namirenje vjerovnika i rehabilitaciju povjerenja? *Ekonomski pregled*, 65(4), 294-317. <https://hrcak.srce.hr/126359>
- Sfakianakis, E. (2021). Bankruptcy prediction model for listed companies in Greece. *Investment Management and Financial Innovations*, 18(2), 166-180. [https://doi.org/10.21511/imfi.18\(2\).2021.14](https://doi.org/10.21511/imfi.18(2).2021.14)
- Šarlija, N. & Jeger, M. (2011). Comparing financial distress prediction models before and during recession. *Croatian Operational Research Review*, 2(1), 133-142. <https://hrcak.srce.hr/96658>
- Vareško, A. & Zubin, C. (2022). Utjecaj pandemije COVID-19 na poslovanje sektora ugostiteljstva u Istarskoj županiji. *Elektornički Zbornik Radova Veleučilišta u Šibeniku*, 16(1-2), 23-37. <https://doi.org/10.51650/ezrvs.16.1-2.2>
- Vehtari, A., Gelman, A. & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413-1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Vuković, B., Milutinović, S., Milićević, N. & Jakšić, D. (2020). Corporate Bankruptcy Prediction: Evidence from Wholesale Companies in the Western European Countries. *Ekonomický Časopis*, 68(5), 477-498. <https://ekonom.sav.sk/en/casopis/rocnik/5-2020>
- Zavrgen, C. V. (1985). Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis. *Journal of Business Finance & Accounting*, 12(1), 19-45. <https://doi.org/10.1111/j.1468-5957.1985.tb00077.x>
- Zenzerović, R. & Peruško, T. (2006). Kratki osvrt na modele za predviđanje stečaja. *Economic Research*, 19(2), 132-151. <https://hrcak.srce.hr/21495>
- Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*, 22, 59-82. <https://doi.org/10.2307/2490859>

A Bayesian approach to logistic regression for bankruptcy prediction of trading companies under uncertainty conditions

Summary

This article provides solutions to common methodological issues in studies dealing with bankruptcy prediction of trading companies, such as the selection of a homogeneous random sample, the choice of financial and non-financial indicators relevant for bankruptcy prediction, determination of an optimal probability threshold for classification, the specification of prior distributions for Bayesian logistic regression applicability, and thus it adds value to the existing literature. Considering the necessity of bankruptcy prediction as the most effective way of its prevention, particularly during crisis periods, requirements for improving forecasting models are imposed. In this context, it was examined not only which indicators are significant for bankruptcy prediction in the wholesale and retail trade sector but also whether the Bayesian approach contributes to the predictive ability of logistic regression during uncertain periods, such as the COVID pandemic when the same method was employed. The adequacy of the Cauchy prior for the constant term is proved, while for the remaining coefficients the Student's t-prior with a standard deviation of 2.5 on the logarithmic odds ratio scale is more applicable. The short-term assets turnover is identified as the indicator that best discriminates bankrupt from non-bankrupt companies, although it is concurrently the most uncertain. Nevertheless, the EBIT margin exhibits the lowest discriminatory power, but compared to the other variables is the most certain, while indebtedness and current liquidity are mod-

erately uncertain variables. The predictive accuracy of Bayesian logistic regression of 90.22% in 2021 is documented by cross-validation, whereby 99.61% of bankrupt trading companies are correctly classified, with a lower cut-off probability compared to standard logistic regression. Adequately specified weakly informative prior's, which are crucial for posterior analysis of the parameters, are the main contribution of the research, and consequently the determination of uncertainty level for each bankruptcy indicator individually.

Keywords: Bayesian logistic regression, bankruptcy prediction, financial and non-financial indicators, wholesale and retail trade, prior distributions of parameters