

PRIMENA KVANTITATIVNIH METODA U PREDVIĐANJU POSLOVANJA PRIVREDNIH DRUŠTAVA

Slobodan Stanojević⁴⁹, Nenad Đorđević⁵⁰, Dejan Volf⁵¹

Pregledni rad
UDK: 005.521:334.7
005.311.12:519.85

Apstrakt

Predikcija poslovanja privrede Srbije je od velike važnosti za investicione aktivnosti, koje su generator rasta i razvoja ekonomije jedne zemlje. Zvanično obelodanjeni podaci o rezultatima poslovanja privrede Srbije su potrebna i dovoljna osnova za sušastvenu analizu finansijskih izveštaja u domenu predikcije, kao i detektovanja potencijalnih gubitaka versus dobitaka u privredi. U prilogu je izvršena analiza ključnih salda računa bilansa uspeha, koji prediktivno detektuju privredna društva na gubitke ili dobitke, primenom vetačke inteligencije tj. mašinskog učenja (Data Mininga). Kvantitativna analiza putem mašinskog učenja se odnosi na bilanse uspeha i stanja, preciznije respektivnih salda računa bilansa uspeha na reprezentativnom uzorku od oko 600 privrednih društava, koja su podvrgnuta analizi sa singnifikantnim rezultatima apsolutne tačnosti.

Ključne reči: finansijski izveštaji, poslovanje,, mašinsko učenje, stablo odlučivanja, selekcija atributa

Uvod

Zvanično obelodanjeno saopštenje o poslovanju privrede u Republici Srbiji za period od 2010. do 2012. godine ukazuju na dramatičan pad ukupne domaće ekonomske aktivnosti praćene smanjenjem industrijske proizvodnje, dalji rast spoljnotrgovinskog deficita, mali rast izvoza, snažnu inflaciju, hronični nedostatak neophodne akumulacije, finansiranje nepokrivene potrošnje i investicija kroz zaduživanje (Koteri, Lester, 2012). Uporedni pokazatelji sa

⁴⁹ Vanredni profesor, dr Slobodan Stanojević, Visoka škola modernog biznisa, Beograd, Terazije br. 27, 11000 Beograd, E-mail: slobodan.stanojevic@mbs.edu.rs

⁵⁰ Nenad Đorđević, SO Đorđević, TC Medaković, Braće Smić br. 23a, Beograd

⁵¹ Dejan Volf, Univerzitet Privredna akademija, Novi Sad, Ulica Cvečarska br. 2

odabranom grupom zemalja još jasnije ukazuju na slabljenje privrede. Naime, Srbija je u grupi zemalja sa najvišom stopom nezaposlenosti, a prednjači u Evropi sa najvećom kumulativnom stopom inflacije i najmanjom stopom ukupnih investicija. Gubici Srpske privrede beleže rast 12.260 od miliona eura u 2010. godini preko projektovanog gubitka do 4.716 miliona eura 2012. godini.

U prilogu ćemo pokušati da na osnovu finansijskog izveštaja (bilans uspeha) za period od 2010. do 2012. godine izvršimo predikciju finansijskih rezultata, preciznije gubitaka nasuprot dobitaka putem detektovanja ključnih salda računa bilansa uspeha.

1. Podaci

Predikciju putem detektovanja ključnih atributa, koji indiciraju gubitak ili dobitak u analizi poslovnih rezultata, kako smo napred napomenuli vršimo primenom mašinskog učenja (Data Mining - DM metoda), a na osnovu odabranog uzorka finansijskih izveštaja: bilans uspeha, i to na relevantnim atributima poslovanja: saldima računa bilansa uspeha za dinamički period od 2010. do 2012. godine. Podaci o elementima finansijskih rezultata saldima računa su korišćeni iz zvanično obelodanjenog Saopštenja o poslovanju privrede Republike Srbije, koje je objavila Republička agencija za privredne registre na svom sajtu (Stanojević, Kaludjerović, Brzaković, Stanojević, 2015).

2. Varijable

Varijable - atributi sa kojima smo u ovoj analizi operisali su iskazani putem salda računa bilansa uspeha kao što su selektirana salda račun bilansa uspeha (Tabela 1.) na osnovu kojih se vrši prediktovanje (predviđjanje).

Tabela 1: Indikatori selektiranih salda računa bilansa uspeha

Atributi selektirana salda bilansa uspeha	
1. Selektirana salda računa bilansa uspeha	12. FINANSIJSKI PRIHODI
2. Prihodi od prodaje	13. FINANSIJSKI RASHODI
3. Prihodi od aktiviranja ucinaka	14. OSTALI PRIHODI
4. Povećanje vrednosti zaliha ucinaka	15. OSTALI RASHODI
	16. NETO DOBITAK POSLOVANJA

5. smanjenje vrednosti zaliha ucinaka 6. Ostali poslovni prihodi 7. Nabavna vredost prodate robe 8. Troškovi materijala 9. Troškovi zarada I naknada 10. Troškovi amortizacije I rezervisanja 11. Ostali poslovni rashodi	KOJE SE OBUSTAVLJA 17. NETO GUBITAK POSLOVANJA KOJE SE OBUSTAVLJA 18. Poreski rashod perioda 19. Odloženi poreski rashodi perioda 20. Odloženi poreski prihodi perioda 21. Isplacena licna primanja poslodavcu 22. Ciljni atribut (Target): dobitaš ili gubitaš
---	---

Privrednih društava kao gubitaša ili dobitaša. U ovom prilogu su primenjena tri najčešće primenjivana metoda Data Mining, i to Stablo odlučivanja (Decision Tree, DT) i metod selekcije atributa (Feature Selection).

3. Primena kvantitativnih metoda (DATA MINING)

Data Mining je kvantitativna metoda iterativnog procesa otkrivanja relacija izmedju atributa, koji su predmet ispitivanja. Suština DM jeste da scenario istraživačkog procesa ne polazi od unapred definisanih hipoteza koje konstituišu “tražene” rezultate. Smatrali smo da je ovakva metodološka konstrukcija domena primene DM od značaja za predikciju determinisanja dobitaša vs gubitaša tj. finansijske klasifikacije na privredne subjekte koji su gubitaši od onih koji nisu. Primenu DM metoda smo usmerili na postupak utvrđivanje određenih salda računa bilansa uspeha koji signifikatno detektuju društva gubitaša od društva dobitaša.

U domenu primene DM identifikovanje dobitaša vs gubitaša u finansijskim izveštajima je tipičan problem klasifikacije, koji sadrži dve procedure. U prvoj, model se uči primenom obučavajućeg uzorka (nadgledano učenje), dok se u drugom vrši klasifikacija na one koji pripadaju determinisanom skupu ili ne pripadaju kao završna faza validacija tačnosti rezultata uzorka.

3.1. Metod učenja stabala odlučivanja (Decision Trees- DT)

Učenje stabala odlučivanja je proces kreiranja diskriminacione funkcije u obliku stabla odlučivanja (Breiman, Friedman, Olshen, Stone, 1984). Stablo se kreira rekursivno, od vrha (korena) prema listovima, tako da svaki čvor stabla predstavlja logički test vrednosti nekog atributa iz opisa problema, a listovi predstavljaju klasu u koju se primer klasifikuje. Prilikom kreiranja, izbor

atributa za svaki čvor vrši se heurističkim metodima, na osnovu ocene kvaliteta diskriminacije (pod)skupa primera iz obučavajućeg skupa, preostalih za diskriminaciju u posmatranom čvoru (leaves). Iako stablo može perfektno da klasifikuje sve primere iz obučavajućeg skupa, to ne predstavlja garanciju visoke tačnosti na novim primerima, jer su često preterano podešena (overfits) prema obučavajućim primerima, pa se vrši pojednostavljenje, čime se dobijaju manja stabla, koja su istovremeno i tačnija i razumljivija. U našoj analizi smo koristili poznati algoritama učenja stabala odlučivanja i to C4.5 (Efstathios, Charalambos, Yannis M, 2007), koji su raspoloživi u okviru sistema WEKA (University of Waikato) (Chercassky, Mučier, 2007) u svrhu selektiranja povezanih atributa. Glavna prednost Stabla odlučivanja jeste da pruža značajan način predstavljanja znanja putem ekstrakcije IF–THEN klasifikacionih pravila (classification rules).

3.2. Metode selekcija atributa (Relief)

Formiranje adekvatnog modela vrši se na osnovu predznanja o problemu i često se svodi na izbor odgovarajućeg skupa atributa. Postojanje irelevantnih i redundantnih (nebitnih, izlišnih) atributa u modelu problema negativno utiče na performanse većine metoda induktivnog učenja, pa se takvi atributi često uklanjaju iz razmatranja postupkom prethodne ili ugrađene selekcije atributa (feature selection). Optimalan skup atributa sadrži sve relevantne attribute, dok se redundantni i irelevantni atributi obično izuzimaju iz razmatranja, iako slabo relevantni redundantni atributi potencijalno sadrže informacije, koje mogu da utiču na poboljšanje klasifikacionih performansi u praksi (Written, Frank, 2005). U prilogu će se neki metodi prethodne selekcije atributa ugrađeni u sistem WEKA (Haussler, 1990) koristiti za dodatnu proveru značajnosti (importance) pojedinih atributa iz modela problema.

4. Rezultati analize

Kao što smo napred napomenuli primenili smo tri alternativna modela. Prvo, stablo odlučivanja primenom WEKE (University of Waikato), Tanagre (Ricco RAKOTOMALALA) i Sipina Research Edition software – učenje stabala C45(Kohavi, 1995). Model je gradjen sa intervalom poverenja (confidence interval) 0,05. Koristili smo ceo uzorak kao obučavajući skup. Dobijeno je stablo odlučivanja, tj. sledeće struktura IF–THEN klasifikacionih pravila.

Model je testirana u odnosu na obučavajući skup i ispravno je klasifikovao dobitaše sa 96% tačnosti i gubitaše sa takodje 94% tačnosti.

Kao što vidimo na Slici 1. Algoritam je iznašao povezujući (splitter) atribut Poreski rashod perioda.

- Poreski rashod perioda < 460,5000
 - Nabavna vredost prodane robe < 5,0896 then KLASA = **DOBITAK** (100,00 % of 26 examples)
 - Nabavna vredost prodane robe >= 5,0896
 - Prihodi od aktiviranja ucinaka < 10552,2217 then KLASA = **GUBITAK** (92,89 % of 211 examples)
 - Prihodi od aktiviranja ucinaka >= 10552,2217
 - Troškovi amortizacije i rezervisanja < 131089,3438 then KLASA = **DOBITAK** (100,00 % of 36 examples)
 - Troškovi amortizacije i rezervisanja >= 131089,3438 then KLASA = **GUBITAK** (89,69 % of 97 examples)
 - Poreski rashod perioda >= 460,5000 then KLASA = **DOBITAK** (93,91 % of 230 examples)
-

Stabla odlučivanja za broj čvorova-leaves (12) sa stopom greške od 4%

U slučaju 12 čvorova dobijeni potkrepljuju rezultate primene metoda selekcije atributa Relief, jer se pojavljuju kao ključni atributi oni koji su tom metodom selektirani (vidi 4.2).

Slika 2: Stablo odlučivanja sa 12 čvorova (leaves)

- Poreski rashod perioda < 460,5000
 - **Nabavna vredost prodane robe** < 5,0900 then KLASA = **DOBITAK** (100,00 % of 26 examples)
 - Nabavna vredost prodane robe >= 5,0900
 - **Prihodi od aktiviranja ucinaka** < 10552,2207
 - OSTALI PRIHODI < 225792,0000
 - FINANSIJSKI PRIHODI < 98286,5000 then KLASA = **GUBITAK** (100,00 % of 157 examples)
 - FINANSIJSKI PRIHODI >= 98286,5000
 - Troškovi zarada I naknada < 15803,7148 then KLASA = **DOBITAK** (57,14 % of 14 examples)

- Troskovi zarada I naknada >= 15803,7148 then **KLASA = GUBITAK** (100,00 % of 22 examples)
 - OSTALI PRIHODI >= 225792,0000 then **KLASA = GUBITAK** (61,11 % of 18 examples)
- **Prihodi od aktiviranja ucinaka >= 10552,2207**
 - Troskovi amortizacije I rezervisanja < 131089,3438 then **KLASA = DOBITAK** (100,00 % of 36 examples)
 - Troskovi amortizacije i rezervisanja >= 131089,3438
 - Prihodi od aktiviranja ucinaka < 40603,3672 then **KLASA = GUBITAK** (100,00 % of 45 examples)
 - **Prihodi od aktiviranja ucinaka >= 40603,3672**
 - **FINANSIJSKI RASHODI < 1339622,0000**
 - **FINANSIJSKI PRIHODI < 205354,5000 then KLASA = GUBITAK** (90,48 % of 21 examples)
 - **FINANSIJSKI PRIHODI >= 205354,5000 then KLASA = DOBITAK** (66,67 % of 12 examples)
 - **FINANSIJSKI RASHODI >= 1339622,0000 then KLASA = GUBITAK** (100,00 % of 19 examples)
- Poreski rashod perioda >= 460,5000
 - smanjenje vrednosti zaliha ucinaka < 105040,0781 then **KLASA = DOBITAK** (97,24 % of 217 examples)
 - smanjenje vrednosti zaliha ucinaka >= 105040,0781 then **KLASA = GUBITAK** (66,67 % of 12 examples)

Preciznije, stablo odlučivanja je identifikovalo sa tačnošću od 96% (za 8 čvorova) i 94%(za 12 čvorova) da je klasifikacija na dobitaše i gubitaše bazirana na poreskim rashodima perioda, čija je struktura IF – THAN opredeljena atributima – saldima računa bilansa uspeha, kada je u pitanju prva i druga iteracija (osim što je druga iteracija sa 12 čvorova razudjenija na više atributa), i to: nabavna vrednost prodane robe koji ako su manji od 5,0896 iniciraju gubitak, ali ako su veći onda uključuju i islov da prihodi učinaka ako su manji od 10552,2217 iniciraju gubitak. Ako su veći od e cifre onda uključuju i islov da troškovi amortizacije i rezervisanja budu veći od 131089,3438 tada iniciraju dobitak da ako su veći od datog iznosa iniciraju gubitak. Sa druge strane ako su poreski rashodi perioda veći od 460,500 inicira

se dobitak čak u 230 slučajeva tj. 93,91%. Poreski prihodi perioda je ključni saldio računa za opredeljenje gubitaka versus dobitaka, uvažavajući da atributi pomenuti u rezultatima su po težini daljeg opredeljenja gubitaka i dobitaka takodje odlučnog značaja. Dalja simulacija sa smanjenjem broja čvorova dovodi do pozicioniranja i atributa povećanja vrednosti zaliha kao značajnog za predviđanje gubitaka vs dobitaka.

4.2. Rezultati primena metoda selekcije atributa

Rezultati WEKA selekcija atributa (2 klase). Različitim metodima pretraživanja i ocenjivanja podskupova atributa, pronalazi se najbolji podskup, koji daje najtačnija pravila (stabla). Neki od metoda za pojedinačne attribute daju i numeričke ocene.

Metod Relieff (ocenjuje se svaki atribut ponaosob), dao je sledeće rezultate:

Selekcija ulaznih atributa

Selektirano od 20 svega 4 ključna Ranked atributa po težinama doprinosa (Weight)

N	Attribute	Weight
1	Nabavna vredost prodane robe	0,021192
2	Smanjenje vrednosti zaliha učinaka	0,008632
3	Prihodi od aktiviranja učinaka	0,007958
4	Povećanje vrednosti zaliha učinaka	0,006025

Metodom Relieff je ocenjeno da su najvažniji atributi hijerarhijski poredjani: Nabavna vredost prodane robe, Smanjenje vrednosti zaliha učinaka, Prihodi od aktiviranja učinaka, Povećanje vrednosti zaliha učinaka, svi ostali atributi imaju

vrednost ispod 0,002 učešća na skali važnosti atributa. Neslaganje u važnosti faktora proističe iz izbora broja čvorova prilikom obrade podataka putem Stabla odlučivanja koje je određeno sa 15. U slučaju kada se taj broj smanji dolazi do izražaja rezultati, koji su dobijeni metodom selekcije atributa.

Zaključak

Primena metoda mašinskog učenja ili pronalaženja utopljenog znanja, ili implicitno znanje (Data Mining) u domenu analize finansijskih izveštaja nije bio čest predmet istraživanja na našim prostorima. Data Mining metode bez sumnje imaju moćne klasifikacione i prediktivne mogućnosti i mogu poslužiti za utvrđivanje potencijalnih gubitaka vs dobitaka, posebno kod donošenja odluke o investicionim aktivnostima.

Cilj našeg istraživanja jeste ispitivanje korisnosti metoda Data Mining u postupku predikcije i detektovanja ključnih salda računa bilansa uspeha javno publikovanih finansijskih izveštaja (bilansa stanja i bilansa uspeha). Primenili smo tri metode metodu Stabla odlučivanja i metodu indukcionih pravila, kao i metodu selekcije atributa.

Rezultati istraživanja potvrđuju naše pretpostavke da javno obelodanjeni finansijski izveštaji sadrže ključne indikatore za gubitke versus dobitke. Sa primenjenim podacima postignuta je velika tačnost preciznije, DM metoda se pokazala dominantnom i sa velikim procentom tačnosti. Naša istraživanja ukazuju da signifikantnost t predikcije u postupku kvalifikacije ključnih atributa – salda računa bilansa uspeha u našem slučaju jeste od 96% do 94% tačnosti. Naravno za potrebe istraživanja moguće je simulovati i manji i veći broj čvorova, ali se broj ključnih faktora povećava, što unekoliko obezvređuje istraživanje. U našem slučaju prediktivna analiza je ukazala da je dominantna atribut poreski rashod perioda, praćen hijerarhijskom signifikantnošću ostalih salda računa bilansa uspeha što pokazuju konkretni rezultati analize selekcije atributa.

Naravno, primena mašinskog učenja nije ograničena samo na predikciju i klasifikaciju, njena suštinska odrednica jeste da stvara precizne uslove za postavku adekvatnih hipoteza, jer metodološki ne podleže unapred traženim rezultatima, koji u značajnoj meri usmeravaju naučne napore, već iz uzorka,

koji može biti i oskudan stvara potpuno signifikatne uslove za postavku pravih, a ne traženih hipoteza, što će biti predmet narednih potpunijih istraživanja.

Literatura

1. Breiman L., Friedman J.H., Olshen R.A., Stone C.J. 1984. Classification and Regression Trees, Wadsworth, Belmont.
2. Chercassky V., Mučier F.M. 2007. Learning from Data: Concept, Theory and Methods, 2ed, Jogn Wiley –IEEE Press.
3. Efstathios K., Charalambos S., Yannis M, 2007, Data Mining techniques for the detection of fraudulent financial statements, Expert System with Application (32): 995-1003.
4. Farlov S. 1984. Self-Organizing methods in Modeling: GMDH type Algorithm, Taylor and Francis.
5. Forsyth R. 1989. Machine Learning: Principles and technics, London: Chapman and Hall.
6. Gilad-Bachrach, Navot, Fishby, „Large margin principles for feature selection“, In Guvon, Gunn, Sikraves, Zadeh. 2006. Feature extraction, foundations and applications, Springer-Verlag.
7. Gilad-Bachrach, Navot, Fishby. 2004. Margin based feature selection - theory and algorithms“, In Proc. 21st ICML.
8. Han J., & Camber, M. 2000. Data mining concepts and techniques, San Diego, USA: Morgan Kaufman.
9. Hart A., 1989. Machine induction as a form of knowledge acquisition in knowledge engineering“, in Forsyth. R, ed. Machine Learning: Principles and techniques, Chapman and Hall, London.
10. Haussler D. 1990. Probably approximately correct learning. In Proc. of the 8 th National Conference on Artificial Intelligence, pp. 1101-1108, Morgan Kaufmann.
11. IMF World Economic Outlook Database, 2013, WB DataBank.
12. Kantardzic M. 2013. Data mining: concepts, models, methods, and algorithms. Wiley-IEEE Press.
13. Kohavi R. 1995. A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection, in Proc. of International Joint Conference on Artificial Intelligence.
14. Koteri S., Lester R. 2012. The Role of Accounting in the Financial Crisis: Lessons For The Future, Accounting Horizons Vol. 26. No.2 2012, pp. 335-352.

15. Quinlan R.J., Cameron-Jonas R.R. 1995. Introduction of Logic Programs:FOIL and Related Systems, in New Generation Computing, (13): 287-312.
16. Stanojević Lj., Kaludjerović N., Brzaković T., Stanojević S. 2015. Detecting Hidden Losses in the Financial Statements – Case of Serbia, International Journal of Business and Social Science, 6(4): 125-132.
17. Written I.H., Frank, E. 2005. Data Mining: Practical machine learning tools and techniques, 2 end edition, Morgan Kaufman, San Francisco.

APPLYING THE MACHINE LEARNING METHOD IN PREDICTING BUSINESS WINNERS VS. LOSERS THROUGH FINANCIAL REPORTS

Slobodan Stanojević⁵², Nenad Đorđević⁵³, Dejan Volf⁵⁴

Abstract

Predicting the management of Serbia's economy is of great importance for investment activities which generate growth and development of a country. Official public data on the results of Serbia's economy management are a necessary and sufficient basis for a fundamental analysis of financial reports in the field of prediction, as well as detecting potential losers vs. winners in the economy. The article performs an analysis of key ratio indicators that detect in a predictive fashion business losers or winners by applying artificial intelligence or machine learning (Data Mining). Quantitative analysis by way of machine learning is applied on balance sheets and income statements, more accurately from a representative sample of about 600 companies which are analyzed with significant results of absolute accuracy.

Keywords: *financial reports, winners vs. losers, prediction, ratio indicators, machine learning*

Rad primljen: 18.09.2016.

Rad prihvaćen: 12.02.2017.

⁵² Slobodan Stanojević Ph.D., Associate Professor, Modern Business School, Terazije no. 27, 11000Belgrade, E-mail: slobodan.stanojevic@mbs.edu.rs

⁵³ Nenad Đorđević, SO Đorđević, Shopping Mall Medaković, Braće Srdić no. 23a, Belgrade.

⁵⁴ Dejan Volf, University Business Accademy, Novi Sad, Cvečarska no. 2