

Metody výběru proměnných pro predikci bankrotu firem

Monika Šebestová

Abstract: Predikce finančního selhání je důležitou činností, prováděnou finančními institucemi za účelem zhodnocení finančního zdraví firem i jednotlivců. V současné době jsou predikční modely založeny na statistických metodách či technikách umělé inteligence. Výběr správných proměnných vstupujících do těchto modelů je důležitým krokem k získání reprezentativního vzorku dat a zvýšení finální predikční přesnosti. Vzhledem k tomu, že neexistuje obecný rámec ukazatelů, pomocí kterých by měl být bankrot podniků predikován, je třeba použít metody, které s tímto výběrem mohou pomoci. Předložený článek se zabývá představením používaných metod a snaží se vyslovit odpověď na otázku, která metoda je pro výběr proměnných nejlepší. Z provedené analýzy vyplynulo, že žádnou metodu pro výběr prvků nelze označit za „nejlepší“ pro predikci bankrotu podniků, protože jejich účinnost značně závisí na použitém predikčním modelu.

Purpose of the article: Předložený článek byl zpracován za účelem zmapování metod, které se používají pro výběr prvků (proměnných) vstupujících do bankrotních modelů. Výběr správných ukazatelů je důležitým krokem k získání reprezentativního vzorku dat a zvýšení finální predikční schopnosti. Proto je důležité zhodnotit, které metody se pro tyto účely hodí nejlépe a zanalyzovat s jakou úspěšností byly použity v dřívějších studiích.

Methodology/methods: Tento článek byl zpracován prostřednictvím analýzy sekundárních dat a předložen ve formě rešerše, která popisuje vývoj metod pro výběr proměnných (finančních ukazatelů) vstupujících do predikčních modelů předpovídajících bankrot podniků.

Scientific aim: Cílem článku je teoretické srovnání metod pro výběr proměnných, používaných v predikci bankrotů firem.

Findings: Z provedené analýzy vyplynulo, že neexistuje univerzální metoda, která by dosahovala nejlepších výsledků při použití jak ve statistických predikčních modelech, tak modelech založených na umělé inteligenci. Vždy je třeba přihlížet k typu predikčního modelu a na jeho základě vybrat nejvhodnější metody pro výběr proměnných.

Conclusions: Lze konstatovat, že žádnou metodu pro výběr prvků nelze označit za „nejlepší“ pro predikci bankrotu podniků. Jejich účinnost značně závisí na použitém predikčním modelu.

Keywords: bankrot, finanční ukazatel, t-test, PCA analýza, faktorová analýza, predikční model

JEL Classification: M15, M21

Úvod

Cílem článku je teoretické srovnání metod pro výběr proměnných, používaných v predikci bankrotů firem. Výběr správných ukazatelů je důležitým krokem k získání reprezentativního vzorku dat a zvýšení finální predikční schopnosti.

Pokrok v oblasti informačních technologií umožňuje získávat informace o rizikových stavech společností mnoha způsoby, jako jsou profesionální agentury, sdělovací prostředky apod. Mnoho věřitelů však obvykle spoléhá na názor analytiků, jejichž úsudek může výsledky analýzy významně ovlivnit. Z tohoto důvodu je vhodnější použít pro identifikaci důležitých prediktivních faktorů statistické metody a metody umělé inteligence.

Mnoho souvisejících studií, které se zabývají tvorbou modelů pro predikci bankrotů firem, nebere v úvahu proces výběru klíčových ukazatelů. Výsledkem jsou potom modely obsahující nadbytečné a redundantní informace, snižující stupeň přesnosti vytvořeného modelu.

Existuje mnoho metod pro výběr proměnných vstupujících do bankrotních modelů. Vybrané metody jsou v předloženém článku představeny a vyslovena odpověď na otázku „*Která metoda je v oblasti predikce finančního selhání podniků nejlepší pro výběr vstupních proměnných?*“

1 Predikce bankrotu

Předpovídání bankrotu firem lze považovat za klasifikační problém. Pomocí vstupních proměnných, tvořených finančními a účetními daty, se bankrotní modely snaží zjistit, zda se podnik řadí do kategorie bankrotujících či nebankrotujících podniků (Tsai, 2009). Predikce bankrotu je důležitá jak pro firmy v jejich rizikovém řízení, tak zejména pro finanční instituce. Finanční instituce potřebují disponovat efektivními prediktivními modely, které jim umožní přijímat vhodná rozhodnutí o poskytování úvěrů (Liang et al., 2015).

V minulosti použil Beaver (1966) finanční ukazatele k predikci zdravých a bankrotujících podniků pomocí lineární regrese. Altman (1968) pro stejné účely použil mnohorozměrnou diskriminační analýzu. V současné době se lze setkat také s modely založenými na technikách data miningu (Geng et al., 2015), které v mnoha případech vykazují lepších výsledků než statistické modely. Zařadit sem lze například neuronové sítě a fuzzy logiku (Mohamed a Salama, 2013).

Mezi další autory, kteří použili některou z metod pro výběr prvků vstupujících do bankrotních modelů lze zařadit Zhou et al. (2015).

2 Metody výběru proměnných pro predikci bankrotu

V mnoha oblastech výzkumu, jako je například modelování systémů a rozeznávání vzorů, je důležité pro vstupy zvolit skupinu atributů s co nejvyšší vypovídací hodnotou. Vzhledem k tomu, že neexistuje obecný rámec finančních ukazatelů, pomocí nichž by měl být bankrot podniku predikován, je důležité z množiny dat vybrat ukazatele reprezentativní a s dostatečnou vysvětlující silou (Liang et al., 2015). Snížení počtu redundantních a irelevantních dat rapidně sníží čas učícího se algoritmu a poskytne obecnější koncept navrženého modelu (Tsai, 2009).

Existuje mnoho známých technik pro extrakci proměnných, jako je například analýza hlavních komponent (PCA), faktorová analýza, diskriminační analýza apod. Výběr proměnných se obvykle provádí před samotným tréninkem modelu. Avšak ne ve všech studiích se lze s touto selekcí proměnných setkat.

Obecně lze metody pro výběr prvků rozdělit do tří skupin, a to filtrovací metody, wrapper metody a hybridní metody (Liang et al., 2015). Filtrovací metody bývají obvykle založeny na statistických technikách a vybírají z určité množiny proměnných podmnožinu, která nejlépe splňuje požadované podmínky. Wrapper metody „obalují“ použitý algoritmus strojového učení, který v modelu slouží k vyhodnocování vybrané podmnožiny atributů. Tento algoritmus je brán jako „černá skříňka“, důležité jsou jen výsledky vyhodnocení. Hybridní metody představují kombinaci předchozích dvou metod s využitím různých kritérií hodnocení.

V posledních letech byly filtrovací a wrapper metody pro výběr proměnných bankrotujících firem použity v mnoha studiích (Liang et al., 2015; Divsalar et al., 2012; Feki et al., 2012). Lin et al. (2014) navrhli integrovaný přístup výběru ukazatelů pro predikci bankrotu firem založený na kombinaci

znalostí odborníků a wrapper metod. Hajek a Michalak (2013) považují wrapper metody za přesnější než metody filtrovací.

Podrobnější popis metod používaných pro predikci bankrotu firem je uveden v následující části.

2.1 Diskriminační analýza

Diskriminační analýza se zabývá klasifikačním problémem a používá se k nalezení lineární kombinace vlastností, které charakterizují nebo oddělují dvě nebo více tříd objektů (Liang, 2015). Klasifikační problém spočívá v pokusu zařadit neznámý objekt do jedné z populací (tříd) na základě vektoru měření provedených na tomto objektu. Tyto populace (třídy) mohou být popsány buď pravděpodobnostním modelem rozložení vektorů měření, nebo od nich existuje reprezentativní vzorek dat (Hendl, 2012).

Diskriminační analýzu ve své studii použili Azayite a Achchab (2016), kteří na základě vybraných ukazatelů sestavili model pro predikci rizika selhání firem v Maroku.

2.2 T-test

Metoda t-test se používá k určení, zda existuje významný rozdíl mezi dvěma skupinami dat. Umožňuje odpovědět na základní otázku: *Pocházejí obě skupiny pozorování ze stejné populace a rozdíly mezi nimi se objevují pouze náhodně nebo mezi těmito skupinami existuje významný rozdíl* (Tsai, 2009)? V predikci bankrotu firem se t-test používá k testování významnosti vlivu finančních ukazatelů na skupinu bankrotujících a nebankrotujících podniků. Pokud mezi nimi existuje významný rozdíl, je finanční ukazatel navržen jako prediktivní proměnná (Zanganeh et al., 2011).

Podle výzkumu Tsai (2009) je metoda t-test nejpřesnější pro výběr ukazatelů vstupujících do modelu predikce bankrotu. K opačnému názoru došli Zanganeh et al. (2011), kteří vyšší přesnosti predikce dosáhli s modelem využívajícím nejčastěji se vyskytující ukazatele v literatuře než s ukazateli vybranými prostřednictvím t-testu.

2.3 ANOVA F-test

Metodu f-testu použili ve své studii Geng et al. (2015), na základě které vybrali ukazatele pro predikční model založený na technikách umělé inteligence. Výběr probíhal na základě síly vztahu mezi každou vstupní proměnnou (finančním ukazatelem) a cílovou proměnnou (bankrotující nebo nebankrotující podnik). Jedná se o analýzu rozptylu, která srovnává několik středních hodnot nezávislých náhodných výběrů. Analýza rozptylu ve své parametrické podobě předpokládá normalitu dat. Pokud tyto podmínky nejsou splněny, je třeba použít neparametrický *Kruskal-Wallisův test*.

Analýzu rozptylu (ANOVA) použili ve svém výzkumu také Fedorova et al. (2013), kteří spolu s ní testovali také další metody pro výběr ukazatelů – vícenásobnou diskriminační analýzu (MDA), klasifikační a regresní stromy a logistickou regresi.

2.4 Korelační matice

Korelační matice zobrazuje korelaci mezi dvěma kvantitativními skupinami dat a současně analyzuje sílu závislosti mezi jejich veličinami. Tato závislost se měří na základě korelačního koeficientu. Vztah mezi dvěma veličinami je vysoce korelovaný, pokud při změně jedné proměnné dochází k podobné změně i druhé veličiny (Tsai, 2009).

2.5 Postupná (stepwise) regrese

K tvorbě modelů založených na regresi existuje metoda, která nalezne optimální podmnožinu prediktorů z větší množiny potenciálních prediktorů. Jedná se o postupnou regresi. Existuje mnoho variant postupné regrese, avšak nejzákladnějším postupem je nalézt jedinou nejlepší proměnnou a postupně přidávat další proměnné, které splňují určité specifické kritérium. Výsledkem je kombinace prediktorů (proměnných), které mají významné korelační koeficienty (Tsai, 2009).

Algoritmus postupné regrese je následující (Hendl, 2012):

- a) V prvním kroku je jako nejlepší prediktor vybrána nezávislá proměnná s největším korelačním koeficientem se závislou proměnnou Y a zařazena do množiny prediktorů.
- b) V dalším kroku se přibere proměnná, která nejlépe zlepšuje predikční schopnost proměnných zařazených do predikce.
- c) Z predikce se odstraní proměnná, jejíž příspěvek pro predikci Y klesl pod určitou úroveň. Pokračuje se krokem b).

Proces přibírání prediktorů (proměnných) končí ve chvíli, kdy žádný další prediktor významně nezlepší predikci (Hendl, 2012).

2.6 Analýza hlavních komponent (PCA)

Hlavní myšlenkou PCA (Principal Component Analysis) je zredukovat počet proměnných, které popisují určitou skupinu proměnných, a sestavit menší počet nových proměnných (komponent), které jsou vzájemně nezávislé. Hlavní komponenty jsou uspořádány podle toho, kolik variability původní množiny proměnných objasňují. Na první pozici potom stojí hlavní komponenta objasňující největší část této variability (Hendl, 2012).

Cílem analýzy je z daného počtu proměnných vytvořit nové proměnné, které jsou nekorelované. Nekorelovanost dat je užitečnou vlastností, protože každá nová proměnná měří jinou vlastnost dat. Nově vzniklé proměnné se nazývají *hlavní komponenty*. Předpokladem použití této analýzy je fakt, že původní data jsou silně korelovaná (Hendl, 2012).

Vhodnost dat pro použití analýzy hlavních komponent je možné ověřit pomocí KMO míry (Kaiser-Meyer-Olkin). Jedná se o statistickou míru, která nabývá hodnot v intervalu $[0;1]$, kde 0 představuje absenci vztahu mezi proměnnými a 1 signalizuje existenci významného vztahu. Předpokladem použití analýzy hlavních komponent je, aby vztahy mezi veličinami byly těsné a nejen zprostředkované vlivem některé ze sousedních proměnných. Doporučená hodnota KMO by měla být minimálně $KMO = 0,6$ a k jejímu výpočtu se používají korelační a parciální koeficienty (Robu a Istrate, 2015).

Do současné doby bylo publikováno několik studií, které se zabývají různými přístupy pro redukci finančních ukazatelů pomocí metody PCA (Tsai, 2009; Achim et al. 2012).

Mezi další autory, kteří se zabývali redukcí počtu ukazatelů metodou PCA, patří Zahedi et al. (2015) a Robu a Istrate (2015).

2.7 Faktorová analýza

Faktorová analýza má podobný cíl jako analýza hlavních komponent, a to popsat chování množiny proměnných pomocí menšího počtu nových proměnných – faktorů. Na základě toho uvádí vzájemné závislosti původních proměnných (Tsai, 2009).

Ve faktorové analýze se více proměnných shlukuje tak, že většina proměnných v jednom shluku spolu silně koreluje. V rámci různých shluků však mezi sebou proměnné korelovat nemají. Faktory jsou interpretovány podle toho, jaké proměnné obsahuje příslušný shluk a variabilita je shrnuta pomocí několika málo faktorů (Hendl, 2012).

Při faktorové analýze se nejdříve určí provizorní *faktorové zátěže*, které udávají, jakou část variability proměnné X vysvětluje faktor F . V druhé fázi se provádí *faktorová rotace*. Představuje transformaci matice faktorových zátěží (Hendl, 2012). Literatura a statistický software nabízí mnoho rotačních algoritmů (např. Varimax, Quartimax,...), jejichž volba závisí na požadované kvalitě interpretace faktorů. Pro lepší interpretaci je dobré, aby faktorové zátěže byly buď blízké 1, nebo 0. Znamená to pak, že korelační koeficient mezi původní proměnnou a faktorem je buď silný, nebo slabý (Analýza hlavních komponent a Faktorová analýza, 2012). Poslední fáze analýzy je tvořena výpočtem *faktorových skóre*, což jsou hodnoty jednotlivých faktorů pro každého jedince nebo změřený objekt.

2.8 Boosted Trees

Karas a Režňáková (2016) použili k ohodnocení statistické významnosti ukazatelů pro bankrotující a nebankrotující podniky metodu Boosted Trees. Tato metoda kombinuje klasifikační a regresní stromy (CART) s boosting algoritmem navrženým J. Friedmanem (2001). S použitím boosting algoritmu se zvyšuje přesnost klasifikačního algoritmu, který je aplikován pro postupné snižování chyby výsledku. Výhodou této neparametrické metody je, že data nemusí pocházet z normálního rozdělení a je schopna zachytit nelineární vztah mezi proměnnými.

2.9 Ostatní metody

Mezi další statistické metody pro výběr prvků lze zařadit *logistickou regresi*, kterou ve svém výzkumu použili Liang et al. (2015). Jedná se o pravděpodobnostní klasifikační model, který pracuje se závisle proměnnou nabývající omezených hodnot (např. binárních). Logistická regrese pak měří vztah mezi kategoričnou závisle proměnnou a jednou nebo více nezávislými proměnnými.

Je třeba poznamenat, že pro výběr reprezentativních prvků modelu lze použít také techniky strojového učení, jako např. *samoorganizační mapy* (SOM) a *genetické algoritmy*, avšak tyto metody jsou v oblasti predikce bankrotu firem používány jen velmi zřídka (Tsai, 2009). Genetické algoritmy se řadí do tzv. obalových (wrapper) technik a jsou založeny na principech genetiky a mechanismech přirozeného výběru. Příbuznou metodou genetických algoritmů je *optimalizace rojem částic* (Particle Swarm Optimization), která vychází z existence určité populace částic, směřujících ke globálnímu optimu (Liang et al., 2013).

Wang (2014) aplikoval na skupinu finančních ukazatelů analýzu *grey relation analysis* (GRA) a zjistil, které ukazatele dostatečně reprezentují model hodnotící výkonnost podniků v Taiwanu. Xu et al. (2014) představili model výběru finančních ukazatelů pomocí „*soft set teorie* (NSS)“.

3 Metody hodnocení

Tsai (2009) vyhodnotil kvalitu výběru proměnných pomocí chyby Typu I a Typu II. Proměnné pro predikci bankrotu firem vybral na základě 5 statistických metod a vytvořil několik modelů založených na vícevrstvých neuronových sítích, které predikovaly finanční selhání firem. Chyba Typu I představuje chybné zařazení nebankrotujících podniků do bankrotujících. Zatímco chyba Typu II řadí bankrotující podniky do nebankrotující skupiny.

Hodnocení kvality predikce podle chyby Typu I a Typu II provedli také Zanganeh et al. (2011), kteří navíc použili také ukazatel predikční přesnosti (Prediction Accuracy). Přesnost predikce lze vypočítat jako procento firem, které jsou správně zařazeny do zdravých nebo bankrotujících firem. Mezi další autory, kteří ověřili přesnost predikci pomocí výše zmíněných metod lze zařadit Liang et al. (2015).

4 Srovnání metod

V literatuře byla popsána řada metod, které umožňují získat proměnné s nejlepší produktivní silou pro navrhovaný model. Není však obecně známo, která z metod se pro tento účel hodí nejlépe, a proto Tsai (2009) porovnal 5 známých metod pro výběr prvků a zjišťoval, která se hodí pro predikci firem nejlépe. Mezi zkoumanými metodami byl t-test, korelační matice, postupná regrese, analýza hlavních komponent (PCA) a faktorová analýza. Zjistil, že nejlepších výsledků bylo dosaženo s použitím metody t-test, která poskytovala nejvyšší predikční přesnost a nejnižší chyby Typu I a Typu II.

Fedorova et al. (2013) použili pro výběr ukazatelů diskriminační analýzu, logistickou regresi a klasifikační a regresní stromy (CRT). Došli k závěru, že nejlepších výsledků dosahoval predikční model založený na neuronových sítích s využitím logistické regrese a CRT. Liang et al. (2013) testovali několik filtračních a wrapper metod pro výběr prvků a došli k závěru, že nejlepších výsledků bylo dosaženo pomocí t-testu, genetických algoritmů a logistické regrese.

Na závěr lze říct, že podle Liang et al. (2013), který testoval metody pro výběr prvků na 4 různých datových souborech, neexistuje nejlepší kombinace metod pro výběr prvků s predikčními modely.

Tabulka 1 porovnává některé související studie z posledních let (2013 - 2017) z hlediska použitých metod výběru prvků a predikčních modelů, na které byla filtrace proměnných aplikována.

Tabulka 16: Srovnání souvisejících studií (Zdroj: vlastní zpracování)

Studie	Metoda výběru	Predikční modely
Wang (2014)	Grey relation analysis	Fuzzy TOPSIS
Fedorova et al. (2013)	Diskriminační analýza Klasifikační a regresní stromy Logistická regrese	Neuronové sítě
Xu et al. (2014)	Soft set teorie (NSS) Analýza hlavních komponent Rough set teorie	Support Vector Machine Neuronové sítě Logistická regrese
Azayite a Achchab (2016)	Diskriminační analýza	Neuronové sítě
Liang et al. (2016)	Diskriminační analýza Logistická regrese T-test Genetické algoritmy	Support Vector Machine K-nearest neighbor Klasifikační a regresní stromy Neuronové sítě
Du Jardin (2017)	Korelační matice	Diskriminační analýza Logistická regrese Neuronové sítě Support Vector Machine

5 Závěr

Predikce bankrotu firem je důležitou součástí finančního rozhodování podniku. Včasné odhalení bankrotujících podniků je již delší dobu diskutovaným tématem, které bylo ve velké míře zkoumáno v účetní a finanční literatuře. V posledních letech jsou hojně využívány k predikci bankrotu techniky data miningu. U těchto technik je důležitým procesem výběr proměnných vstupujících do modelu, jejichž přítomnost zvýší prediktivní schopnost navržených modelů. S tím souvisí filtrace zbytečných a irelevantních informací (vstupů) a současné zvýšení výkonnosti modelu při snížení náročnosti tréninku modelů.

Tento článek se zabýval teoretickým srovnáním metod pro výběr proměnných, používaných v predikci bankrotů firem. Závěrem lze konstatovat, že nelze žádnou metodu pro výběr prvků označit za „nejlepší“ pro predikci bankrotu podniků. Jejich účinnost značně závisí na použitém predikčním modelu. Z analýzy vyplynulo, že jiné metody pro výběr prvků se hodí pro statistické modely a jiné pro modely založené na umělé inteligenci. Pro další výzkum by bylo vhodné zanalyzovat metody výběru proměnných pro konkrétní predikční model založený např. pouze na neuronových sítích nebo fuzzy logice.

Acknowledgment

Příspěvek je výstupem projektu specifického výzkumu FP-J-17-4174 *Využití umělé inteligence v podnikatelství* Interní grantové agentury Vysokého učení technického v Brně.

Použité zdroje

ACHIM, Monica Violeta, Codruta MARE a Sorin Nicolae BORLEA, 2012. *A Statistical Model of Financial Risk Bankruptcy Applied for Romanian Manufacturing Industry*. *Procedia Economics and Finance*. 3, 132-137. DOI: 10.1016/S2212-5671(12)00131-1. ISSN 22125671.

ALTMAN, Edward I, 1968. *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*. *The Journal of Finance*. 23(4), 589-609.

A.MOHAMED, Ahmed a Ahmed S. SALAMA, 2013. *A Fuzzy Logic based Model for Predicting Commercial Banks Financial Failure*. *International Journal of Computer Applications*. 79(11), 16-21. DOI: 10.5120/13785-1817. ISSN 09758887.

Analýza hlavních komponent a Faktorová analýza [online], 2012. [cit. 2017-11-20]. Dostupné z: http://www.fsps.muni.cz/~sebera/vicerozmerna_statistika/pca.html

AZAYITE, Fatima Zahra a Said ACHCHAB, 2016. *Hybrid Discriminant Neural Networks for Bankruptcy Prediction and Risk Scoring*. *Procedia Computer Science*. 83, 670-674. DOI: 10.1016/j.procs.2016.04.149. ISSN 18770509.

BEAVER, Wh., 1966. *Financial Ratios as Predictors of Failure*. *Journal of Accounting Research*. (4), 71-111.

DIVSALAR, Mehdi, Habib ROODSAZ, Farshad VAHDATINIA, Ghassem NOROUZZADEH a Amir Hossein BEHROOZ, 2012. *A Robust Data-Mining Approach to Bankruptcy Prediction*. *Journal of Forecasting*. 31(6), 504-523. DOI: 10.1002/for.1232. ISSN 02776693.

- DU JARDIN, Philippe, 2017. *Dynamics of firm financial evolution and bankruptcy prediction*. Expert Systems with Applications. 75, 25-43. DOI: 10.1016/j.eswa.2017.01.016. ISSN 09574174.
- FEDOROVA, Elena, Evgenii GILENKO a Sergey DOVZHENKO, 2013. *Bankruptcy prediction for Russian companies: Application of combined classifiers*. Expert Systems with Applications. 40(18), 7285-7293. DOI: 10.1016/j.eswa.2013.07.032. ISSN 09574174.
- FEKI, Asma, Anis Ben ISHAK a Saber FEKI, 2012. *Feature selection using Bayesian and multiclass Support Vector Machines approaches: Application to bank risk prediction*. Expert Systems with Applications. 39(3), 3087-3099. DOI: 10.1016/j.eswa.2011.08.172. ISSN 09574174.
- FRIEDMAN, J. H., 2001. *Greedy function approximation: A gradient boosting machine*. Annals of Statistics. 29(5), 1189-1232.
- GENG, Ruibin, Indranil BOSE a Xi CHEN, 2015. *Prediction of financial distress: An empirical study of listed Chinese companies using data mining*. European Journal of Operational Research. 241(1), 236-247. DOI: 10.1016/j.ejor.2014.08.016. ISSN 03772217.
- HAJEK, Petr a Krzysztof MICHALAK, 2013. *Feature selection in corporate credit rating prediction*. Knowledge-Based Systems. 51, 72-84. DOI: 10.1016/j.knosys.2013.07.008. ISSN 09507051.
- HENDL, Jan, 2012. *Přehled statistických metod: analýza a metaanalýza dat*. 4., rozš. vyd. Praha: Portál. ISBN 978-80-262-0200-4.
- KARAS, Michal a Režňáková MÁRIA, 2016. *Bankruptcy prediction: To what degree does past development count?* WSEAS TRANSACTIONS ON BUSINESS AND ECONOMICS. 644-652.
- LIANG, Deron, Chia-Chi LU, Chih-Fong TSAI a Guan-An SHIH, 2016. *Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study*. European Journal of Operational Research. 252(2), 561-572. DOI: 10.1016/j.ejor.2016.01.012. ISSN 03772217.
- LIANG, Deron, Chih-Fong TSAI a Hsin-Ting WU, 2015. *The effect of feature selection on financial distress prediction*. Knowledge-Based Systems. 73, 289-297. DOI: 10.1016/j.knosys.2014.10.010. ISSN 09507051.
- LIN, Fengyi, Deron LIANG, Ching-Chiang YEH a Jui-Chieh HUANG, 2014. *Novel feature selection methods to financial distress prediction*. Expert Systems with Applications. 41(5), 2472-2483. DOI: 10.1016/j.eswa.2013.09.047. ISSN 09574174.
- ROBU, Ioan Bogdan a Costel ISTRATE, 2015. *The Analysis of the Principal Components of the Financial Reporting in the Case of Romanian Listed Companies*. Procedia Economics and Finance. 20, 553-561. DOI: 10.1016/S2212-5671(15)00108-2. ISSN 22125671.
- TSAI, Chih-Fong, 2009. *Feature selection in bankruptcy prediction*. Knowledge-Based Systems. 22(2), 120-127. DOI: 10.1016/j.knosys.2008.08.002. ISSN 09507051.
- WANG, Yu-Jie, 2014. *The evaluation of financial performance for Taiwan container shipping companies by fuzzy TOPSIS*. Applied Soft Computing. 22, 28-35. DOI: 10.1016/j.asoc.2014.03.021. ISSN 15684946.
- XU, Wei, Zhi XIAO, Xin DANG, Daoli YANG a Xianglei YANG, 2014. *Financial ratio selection for business failure prediction using soft set theory*. Knowledge-Based Systems. 63, 59-67. DOI: 10.1016/j.knosys.2014.03.007. ISSN 09507051.

ZAHEDI, Javad a Mohammad Mahdi ROUNAGHI, 2015. *Application of artificial neural network models and principal component analysis method in predicting stock prices on Tehran Stock Exchange*. Physica A: Statistical Mechanics and its Applications. 438, 178-187. DOI: 10.1016/j.physa.2015.06.033. ISSN 03784371.

ZANGANEH, Tayebeh, Meysam RABIEE a Masoud ZAREI, 2011. *Applying Adaptive Neuro - Fuzzy Model for Bankruptcy Prediction*. International Journal of Computer Applications. 20(3), 15-21.

ZHOU, Ligang, Dong LU a Hamido FUJITA, 2015. *The performance of corporate financial distress prediction models with features selection guided by domain knowledge and data mining approaches*. Knowledge-Based Systems. 85, 52-61. DOI: 10.1016/j.knosys.2015.04.017. ISSN 09507051.

Autor(s) contact(s)

Ing. Monika Šebestová
Vysoké učení technické v Brně
Fakulta podnikatelská
Ústav informatiky
Kolejní 2906/4, 612 00 Brno
Česká republika
Tel.: +420 54114 3709
E-mail: sebestova@fbm.vutbr.cz