

BANKROTU DATU ANALĪZES VEIKŠANAS IESPĒJAS AR LAIKRINDU KLASTERIZĀCIJAS PALĪDZĪBU

POSSIBILITIES OF PERFORMING BANKRUPTCY DATA ANALYSIS USING TIME SERIES CLUSTERING

Pēteris GRABUSTS

Dr. sc. ing., asoc. profesors
Rēzeknes Augstskola
Atbrīvošanas al. 90, Rēzekne, LV-4601, Latvija
tālr.: +371 26593165; e-pasts: peter@ru.lv

Abstract. Prediction of corporate bankruptcy is a study topic of great interest. Under the conditions of the modern free market, early diagnostics of unfavourable development trends of company's activity or bankruptcy becomes a matter of great importance. There is no general method which would allow one to forecast unfavourable consequence with a high confidence degree. This paper focuses on the analysis of the approaches that can be used to perform an early bankruptcy diagnostics- in previous research multivariate discriminant analysis (MDA), neural network based approach and rule extraction method have been examined. Lately, time series clustering approach has become popular and its feasibility for bankruptcy data analysis is being investigated. Experiments carried out validate the use of such methods in the given class of tasks. As a novelty, an attempt to apply time series clustering method to the analysis of bankruptcy data is made.

Keywords: bankruptcy prediction, financial ratio, time series, clustering

1. Ievads

Maksātnešpēja vai bankrots ir ar tiesas lēmumu atzīts vai uzņēmuma izsludināts paziņojums par savu nespēju pilnā mērā veikt savas saistības pret kreditoriem un/vai nespēju veikt obligātos maksājumus.

Bankrotu diagnostika jeb prognozēšana kā patstāvīga disciplīna parādījās uzreiz pēc Otrā pasaules kara. To sekmēja straujais bankrotu skaits pēc pasūtījumu samazināšanās kara vajadzībām. Radās nepieciešamība izdalīt svarīgākos finanšu darbības rādītājus, atrast tajos tās likumsakarības, kas noved vai var novest līdz bankrotam.

Sākotnēji bankrotu prognozēšanas jautājumi tika pētīti empīriskā līmenī, kas bieži vien radīja nekorektus secinājumus. Bankrotējošo un nebankrotējošo uzņēmumu rādītāji tika salīdzināti savā starpā un tika dots pieņēmums, kurš rādītājs labāks vai sliktāks. Pirmie nopietnie pētījumi izveidot efektīvu bankrotu prognozēšanas metodiku tika izstrādāti 20. gs. 60. gados un tos sekmēja datortehnikas attīstība.

Vispārpieņemtas ir divas galvenās pieejas bankrotu prognozēšanā. Pirmā pieeja pamatojas uz finanšu datiem un ietver darbības ar dažādiem rādītājiem: ļoti lielu popularitāti ieguvušo Altmana Z-skaitli vai Taflera koeficientiem, kā arī “spēju lasīt bilanci”. Otrā pieeja pamatojas uz datiem par bankrotējušiem uzņēmumiem, kurus salīdzina ar pētāmo uzņēmumu datiem.

Pirmā pieeja tiek uzskatīta par ļoti efektīvu, taču tiek atzīmēti arī būtiski trūkumi:

- uzņēmumi, kam ir finanšu problēmas, kavējas ar savu pārskatu publicēšanu un tādējādi konkrētie dati var būt gadiem ilgi nepieejami;
- ja dati tomēr tiek publiskoti, tie var tikt mākslīgi uzskaistināti, kas arī nesniedz objektīvu informāciju;
- atsevišķi finanšu darbības rādītāji var liecināt par gaidāmo bankrotu, tai pašā laikā citi var dot pamatu uzskatīt uzņēmumu par stabilu. Tādos apstākļos grūti spriest par reālo situāciju.

Otrā pieeja balstās uz jau bankrotējošo uzņēmumu finanšu rādītāju salīdzināšanu ar aizdomīga uzņēmuma tādiem pašiem rādītājiem. Pēdējos gados publikācijās parādīties ļoti daudz bankrotējošo uzņēmumu finanšu rādītāju, taču lielākā daļa nedod priekšstatu par atsevišķu rādītāju svarīguma pakāpi, to prioritātēm un secību. Būtiska nozīme ir arī bankrotējošo uzņēmumu darbības rādītājiem pa gadiem. Daudzi pētnieki izmanto bankrotējošo uzņēmumu finanšu rādītājus “1 gadu pirms bankrota”, “2 gadus pirms bankrota”, kā arī “3 gadus pirms bankrota”.

Bankrotu prognozēšanā tiek izmantotas vairākas metožu grupas, kas ļauj noteikt potenciāla bankrota draudus:

- empīriskā pieeja;
- mākslīgo neironu tīklu pieeja;
- likumu iegūšanas pieeja;
- citas pieejas (ģenētiskie algoritmi, tēlu atpazīšanas metodes).

Turpmākajā izklāstā tiek analizētas atsevišķas metodes bankrotu prognozēšanā - empīriskā pieeja, kā arī neironu tīklu pieeja, kas kļuva ārkārtīgi populāra pagājušā gadsimta 90. gados. Eksperimentālās daļas mērķis - salīdzināt vairākas metodes bankrotu datu analīzes nolūkā un izvērtēt to iespējas. Eksperimenta mērķiem tika izmantoti 63 uzņēmumu bilances dati (46 - bankrotējuši un 17 - nebankrotējuši). No pieejamajiem datiem tika izskaitļoti vairāki finanšu rādītāji, kas tika izmantoti veicamajos eksperimentos. Eksperimentu mērķis bija parādīt, ka ar intelektuālās datu analīzes metodēm ir iespējams analizēt bankrotu datus ar nolūku iegūt secinājumus par potenciālajiem bankrota draudiem.

Ļoti perspektīva var izrādīties finanšu rādītāju traktēšana ar laikrindu palīdzību ar turpmāko to analīzi, izmantojot laikrindu līdzības kritērijus un klasteranalīzi. Kā novitāti var uzrādīt izstrādāto metodiku šāda tipa uzdevumu veikšanai un izpētīto perspektīvo laikrindu klasterizācijas izmantošanu bankrotu datu analīzē.

Bankrota iestāšanās iespējamību ar lielu varbūtības pakāpi ļauj noteikt mūsdienu populārākās prognozēšanas metodes. Tādējādi var izvēlēties optimālāko maksātspējas prognozēšanas metodi.

2. Bankrotu datu finanšu rādītāji un empīriskā pieeja

Bankrota diagnostika ir mērķa finanšu analīzes sistēma, kuras virziens ir krīzes situācijas kontrolēšana uzņēmumos. Šim nolūkam kopējā uzņēmuma finanšu stāvokļa analīzē tiek izveidota atsevišķa finanšu rādītāju grupa, pēc kuriem varētu spriest par bankrota draudiem. Vispārīgā gadījumā nav teorētiska pamatojuma, kādi finanšu rādītāji var tikt lietoti dažādos bankrotu pētīšanas modeļos. Vairāki pētnieki, veicot bankrotējošo uzņēmumu analīzi, aprēķina daudzus rādītājus un pēc tam atlasa potenciāli būtiskākos. Piemēram, Altmana modeli (2.) Z-skaitlis operē ar 5 finanšu rādītājiem. Altmans uzskata, ka šiem rādītājiem ir vislielākā prognozēšanas iespēja:

$X_1 = \text{Tīrais apgrozāmais kapitāls} / \text{Aktīvu kopsumma.}$

$X_2 = \text{Nesadalītā peļņa} / \text{Aktīvu kopsumma.}$

$X_3 = \text{Peļņa pirms nodokļiem} / \text{Aktīvu kopsumma.}$

$X_4 = \text{Akciju tirgus vērtība} / \text{Aizņemtā kapitāls.}$

$X_5 = \text{Neto apgrozījums} / \text{Aktīvu kopsumma.}$

Akadēmiskajos pētījumos ļoti lietderīgi ir dati par 31 finanšu rādītājiem (4.). Neapšaubāmi, ka var pielietot arī citus finanšu rādītājus, kas varētu tikt izmantoti efektīvākai bankrotu prognozēšanai.

Par empīriskās pieejas pionieriem tiek uzskatīti Bīvers (5.) un E. Altmans (2.). Bīvers bija viens no pirmajiem, kas bankrotu pētniecībā pielietoja bilances datus. Viņa analīze bija samērā vienkārša un balstījās uz viena finanšu rādītāja izpēti, salīdzinot to ar citiem rādītājiem. Viņš secināja, ka finanšu rādītājs "naudas plūsma / saistību kopsumma" ir ļoti būtisks faktors, kas jāņem vērā bankrotu analīzē. Bīvera darbi bija iesākums daudzkritēriju analīzes izmantošanai, ko tālāk attīstīja Altmans un citi.

Altmans savos pētījumos izmantoja diskriminantanalīzi (MDA). Altmana Z-skaitlis ir ļoti populārs bankrotu analīzes pētījumos un satur vairākus rādītājus, ko apvieno tā saucamajā Z-skaitlī, pirms tam tos reizinot ar noteiktu korelācijas koeficientu, pēc kuru lieluma var spriest par konkrētā rādītāja nozīmi. Z - skaitli izsaka ar formulu:

$$Z = 1,2 * X1 + 1,4 * X2 + 3,3 * X3 + 0,6 * X4 + X5. \quad (1)$$

- Ja $Z > 3$, tad bankrota iespējamība ir zema un nav nepieciešams veikt sīkāku finanšu stāvokļa analīzi;
- Ja $2,7 < Z \leq 3$, tad bankrots var iestāties. Uzņēmumam ir radušās zināmas problēmas ar maksāspēju, bet uzņēmums tās varētu novērst;
- Ja $1,8 < Z \leq 2,7$, tad bankrota iespējamība ir augsta. Uzņēmumam ir ļoti nopietnas finansiālas problēmas. Nepieciešama rūpīga uzņēmuma finanšu stāvokļa analīze;
- Ja $0 < Z \leq 1,8$, tad bankrota iespējamība ir ļoti augsta. Uzņēmuma finanšu stāvokli var uzlabot tikai radikālas pārmaiņas uzņēmuma finanšu un investīciju jomā.

Minēto rādītāju aprēķins ļauj uzņēmumu vadītājiem novērtēt savu darbību, finanšu rādītājus un laicīgi reaģēt uz esošajām problēmām.

Altmana Z-skaitļi ir populāri ar to, ka praksē tos plaši pielieto kā ieejas datus dažādos prognozēšanas algoritmos.

3. Iepriekšējie pētījumi bankrotu datu analizē

Pētījumi tika veikti ar mērķi salīdzināt vairākas metodes bankrotu datu analīzes nolūkā un izvērtēt to iespējas (6.). Uzņēmumu bankrotu dati ņemti no (10.) un (14.). Pētījuma mērķiem tika izmantoti 63 uzņēmumu bilances dati (46 - bankrotējuši un 17 - nebankrotējuši). Tika pieņemts lēmums no esošajiem datiem izskaitļot šādus finanšu rādītājus un turpmāk tos izmantot visos veicamajos eksperimentos:

- R2: Naudas plūsma / Īstermiņa saistības;
- R3: Naudas plūsma / Aktīvu kopsumma;
- R7: Apgrozāmie līdzekļi / Īstermiņa saistības;
- R9: Apgrozāmie līdzekļi / Aktīvu kopsumma;
- R31: Tīrais apgrozāmais kapitāls / Aktīvu kopsumma.

3.1. Empīriskā pieeja

Empīriskā pieeja demonstrēta uz diskriminantanalīzes MDA pamata, kas tika veikta ar statistiskās paketes SPSS palīdzību. Pielietojuma rezultāti parādīti 1.tabulā.

Rezultātā 82,5% no visiem elementiem ir korekti klasificēti. Neklasificēto elementu numuri ir: 14, 26, 28, 35, 36, 37, 41, 58, 59, 60, 62 (tie ir uzņēmumu dati, kuriem nevar noteikt - ir bankrota stāvoklis vai nav).

Diskriminantanalizes klasifikācijas rezultāti

Bankrots		Prognozējamās grupas elementi		Kopā
		0	1	
Oriģināls	Skaitis	0	39	46
		1	4	17
	%	0	84,8	100
		1	23,5	100

Avots: autora pētījumu rezultāts (6.)

3.2. Mākslīgo neironu tīklu pieeja

Neironu tīklu pieeja bankrotu prognozēšanas jomā kļuva ārkārtīgi populāra pagājušā gadsimta 90. Gados. Šajā laikā tika veikti pētījumi par neironu tīklu izmantošanu bankrotu analizē. Tos veica Odoms un Šarda (9.), Tams un Kiangs (11.) un daudzi citi (3.), (4.).

Vieni no pirmajiem, kas pielietoja neironu tīklu metodes bankrotu analīzes pētījumos, bija Odoms un Šarda. Tīkla ieejā tika izmantoti Altmana Z-skaitļi par 128 uzņēmumiem un tika parādīts, ka neironu tīklu pieeja dod labākus rezultātus par MDA. Tams un Kiangs salīdzināja vairākas bankrotu diagnostikā izmantojamās metodes (MDA, vienslāņu tīklu, daudzslāņu tīklu) un pierādīja, ka datus “1 gads līdz bankrotam” daudzslāņu tīkls ir visefektīvākais (11.).

Analizējot neironu tīklu izmantošanu bankrotu analizē, tika izdarīti šādi secinājumi:

- neironu tīkliem ir apmēram 90% precizitāte, salīdzinot ar 80-85% citām metodēm (MDA);
- bankrotus var paredzēt vairākus gadus līdz to notikšanai, turklāt paredzēšanas precizitāte datiem “2 gadi līdz bankrotam” praktiski neatšķiras no datiem “1 gads līdz bankrotam”.

Bankrotu prognozēšanas uzdevums faktiski ne ar ko neatšķiras no klasiskā tēlu atpazīšanas uzdevuma. Tīkla ieejā tiek doti finanšu rādītāji R, notiek tīkla apmācība, izmantojot slēptos neironus. Tīkla izejā tikai 2 stāvokļi - bankrots un ne bankrots.

Analizējot literatūru, tika secināts, ka pētnieki pārsvarā izmanto daudzslāņu tīklu ar kļūdas atpakaļ izplatības apmācību, tāpēc par neironu tīkla modeli tika izvēlēts tieši šāds tīkls (3.). Eksperimentāli tika konstatēts, ka apmācības nolūkiem piemērotākais ir neironu tīkls ar 5 slēptajiem neironiem, un turpmākie eksperimenti balstījās uz modeli 5-5-1. Izvēlēta tīkla modeļa svarīgākie parametri ir aktivizācijas koeficients α un aktivizācijas funkcijas \tanh slīpums β .

Pirmajā eksperimenta daļā tika pārbaudīta apmācības kvalitātes atkarība no α vērtības. Dotais eksperimentālais modelis parādīja, ka pie α vērtībām robežās līdz 0,8 algoritms vai nu nekonverģē, vai arī nenotiek tikla korekta apmācība. Ticami apmācību rezultāti tika iegūti pie $\alpha=0,8$ un $\alpha=0,9$.

Otrajā eksperimenta daļā tika pārbaudīta apmācības kvalitātes atkarība no aktivizācijas funkcijas \tanh slīpuma $\beta \in [0,1; 1]$. Eksperimentu rezultātā izrādījās, ka tikai pie atsevišķām β vērtībām tika iegūta pieņemama apmācības kvalitāte. 2. tabulā parādīti būtiskākie rezultāti un neklasificētie elementi pie dažādām α un β vērtībām.

2. tabula

Eksperimenta rezultāti (parametri α , β un neklasificētie elementi)

Parametrs α	Parametrs β	Metode	Cikli	Neklasificēto elementu skaits	Neklasificētie elementi
$\alpha = 8,0$	$\beta = 8,0$	NN-1	41	9	14,26,36,37,41,58,59,60,62
	$\beta = 9,0$	NN-2	889	6	37,50,58,59,60,62
$\alpha = 9,0$	$\beta = 8,0$	NN-3	46	9	14,26,35,36,37,41,58,59,62
	$\beta = 9,0$	NN-4	1489	7	37,50,58,59,60,62,63

Avots: autora pētījumu rezultāts (6.)

2. tabulā apkopoti dati par pielietoto metožu neklasificētajiem elementiem, t.i., tiem uzņēmumiem no datu izlases, kurus dotās metodes nespēj korekti klasificēt.

3. tabula

Kopsavilkuma dati par izmantotajām metodēm un to neklasificētajiem elementiem

Metode	Parametri	Neklasificētie elementi												
MDA		14	26	28	35	36	37	41		58	59	60	62	
NN-1	$\alpha = 0,8,$ $\beta = 0,8$	14	26			36	37	41		58	59	60	62	
NN-2	$\alpha = 0,8,$ $\beta = 0,9$						37		50	58	59	60	62	
NN-3	$\alpha = 0,9,$ $\beta = 0,8$	14	26		35	36	37	41		58	59		62	
NN-4	$\alpha = 0,9,$ $\beta = 0,9$						37		50	58	59	60	62	63

Avots: autora pētījumu rezultāts (6.)

No 3. tabulas datiem redzams, ka dotajai bankrota datu izlasei visas metodes nespēj klasificēt 37, 58, 59 un 62 datu elementus. Rēķinot absolūtajos skaitļos, ir redzams, ka NN-2 metode korekti klasificē 90,5% elementu, NN-4 metode - 89% elementu, NN-1 un NN-3 metodes - 85,7%, MDA - 82,5%. Var konstatēt, ka dotajai datu izlasei neironu tīklu metode efektīvāk veic bankrotu datu diagnostiku, kas faktiski atbilst agrāk izteiktajiem secinājumiem par Tama un Kianga rezultātiem.

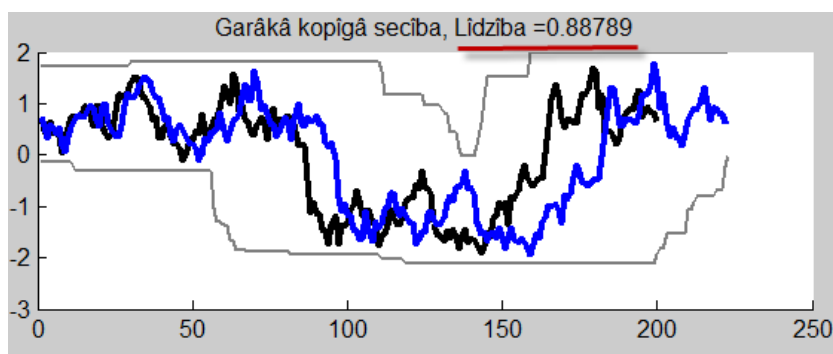
3.3. Laikrindu līdzības novērtēšanas iespējas

Vispārīgā gadījumā laikrinda tiek traktēta kā datu virkne noteiktā laika intervālā (8.), (12.). Laikrindu analīze ir pietiekami labi pazīstams uzdevums, taču pēdējā laikā tiek veikti pētījumi ar nolūku mēģināt pielietot klasterizāciju laikrindu analīzē – laikrindu datu sadalīšanu līdzīgās grupās. Laikrindu attēlošanas klasteru formā galvenā motivācija ir laikrindu raksturlielumu labāka izpratne.

Darba uzdevums bija izpētīt laikrindu klasterizācijas izmantošanas metodoloģiju, apzināt laikrindu tuvības jeb līdzības novērtējumus un izmantot tos laikrindu klasterizācijas rezultātu analīzē. Par līdzības novērtējumu tradicionāli tiek izmantota Eiklīda distance, taču bieži līdzības mēra noteikšanai izmanto vairāk komplicētas metodes, piemēram, garākās kopīgās virknes metodi- LCSS (1.). Laikrindu A un B līdzību izsaka ar skaitli $LCSS(A,B) / \max(|A|, |B|)$, kur $LCSS[i, j]$ nosaka garāko kopīgo secību starp pirmajiem i laikrindas A elementiem un pirmajiem j elementiem no laikrindas B:

$$LCSS[i, j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ 1 + LCSS[i-1, j-1] & \text{if } a_i = b_j \\ \max(LCSS[i-1, j], LCSS[i, j-1]) & \text{otherwise} \end{cases} \quad (2)$$

Tipisks divu laikrindu līdzības noteikšanas ar LCSS metodi piemērs parādīts 1. attēlā.



1. attēls. LCSS metodes demonstrācija

Avots: autora pētījumu rezultāts

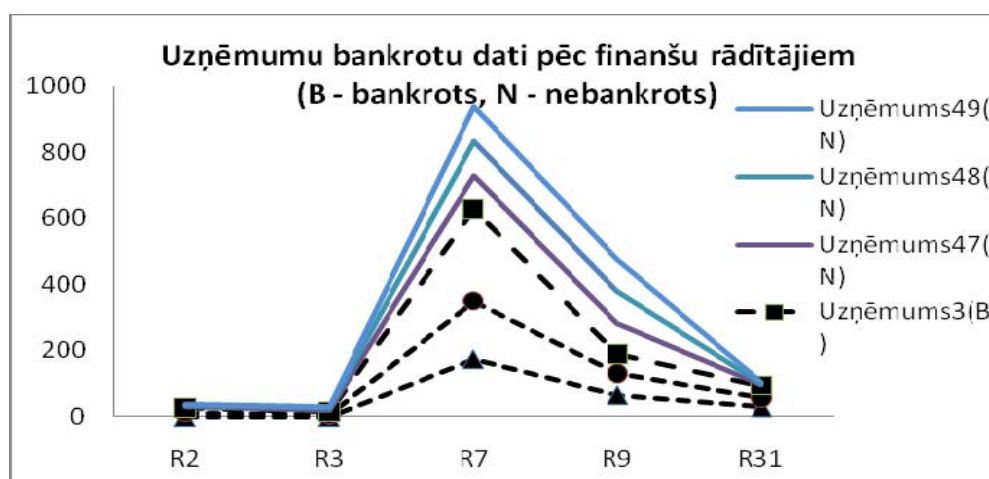
Dotajā etapā tika veikti divi uzdevumi. Pirmais bija noteikt laikrindu līdzības novērtējuma iespējas. Eksperimenta mērķis bija pārbaudīt hipotēzi par LCSS metodes piemērotību bankrotu datu kā laikrindu līdzības novērtēšanai. Tika ņemti 3 bankrotējušu un 3 nebankrotējušu uzņēmumu dati, kas parādīti 4. tabulā un 2. attēlā.

4. tabula

Uzņēmumu finanšu rādītāju dati

Uzņēmums	Finanšu rādītāji					Statuss
	R2	R3	R7	R9	R31	
Uzņēmums1(B)	0	0	175	67	29	Bankrots
Uzņēmums2(B)	13	8	175	65	28	Bankrots
Uzņēmums3(B)	14	8	277	58	37	Bankrots
Uzņēmums47(N)	4	4	101	93	1	Nebankrots
Uzņēmums48(N)	5	5	104	97	4	Nebankrots
Uzņēmums49(N)	4	4	105	95	5	Nebankrots

Avots: adaptēts no (14.)



2. attēls. Uzņēmumu bankrotu dati

Avots: autora pētījumu rezultāts

5. tabulā parādīti LCSS metodes pielietošanas rezultāti - pa pāriem tika salīdzinātas laikkrindas un iegūtas to līdzības vērtības.

5. tabula

Laikrindu pa pāriem salīdzinājuma līdzības vērtības

	Uzņ. 1 (B)	Uzņ. 2 (B)	Uzņ. 3 (B)	Uzņ. 47 (N)	Uzņ. 48 (N)	Uzņ. 49 (N)
Uzņēmums1 (B)	1	0,4	0,4			
Uzņēmums2 (B)	0,4	1	0,4			
Uzņēmums3 (B)	0,4	0,4	1			
Uzņēmums47 (N)				1	0,5	0,4
Uzņēmums48 (N)				0,5	1	0,8
Uzņēmums49 (N)				0,4	0,8	1

Avots: autora pētījumu rezultāts

No 5. tabulas var secināt, ka šajā gadījumā laikrinda Uzņēmums1 ir nedaudz līdzīga laikrindai Uzņēmums2 (0,4) un Uzņēmums3 (0,4). Laikrinda Uzņēmums47 nedaudz līdzīga laikrindai Uzņēmums48 (0,5) un laikrindai Uzņēmums49 (0,4). Laikrindai Uzņēmums48 ir līdzība ar laikrindu Uzņēmums49 (0,8). Var redzēt, ka nevienai no bankrotējušo uzņēmumu laikrindām nav līdzības ar nebankrotējušo uzņēmumu laikrindām.

Analizējot 5. tabulas datus, varētu pieņemt, ka laikrindas Uzņēmums1, Uzņēmums2 un Uzņēmums3 atrodas vienā grupā jeb klasterī, bet laikrindas Uzņēmums47, Uzņēmums48 un Uzņēmums49 - citā klasterī.

Otrais uzdevums bija izpētīt klasterizācijas algoritmu iespējas laikrindu klasterizācijā. Tika izmantots klasiskais k-vidējais klasterizācijas algoritms (7.). Vispārīgā gadījumā klasteranalīzes metodes dod iespēju sadalīt pētāmos objektus līdzīgu objektu grupās, ko sauc par klasteriem. K-vidējais klasterizācijas algoritms minimizē kvalitātes rādītāju, kurš noteikts kā visu punktu, kas pieder klastera apgabalam, attālumu līdz klastera centra kvadrātu summa. Šī procedūra ieguva tādu nosaukumu, pamatojoties uz klasteru grupas iekšienē vidējo attālumu aprēķināšanu līdz klastera centram.

Algoritms k-vidējais izpildās šādu soļu secībā:

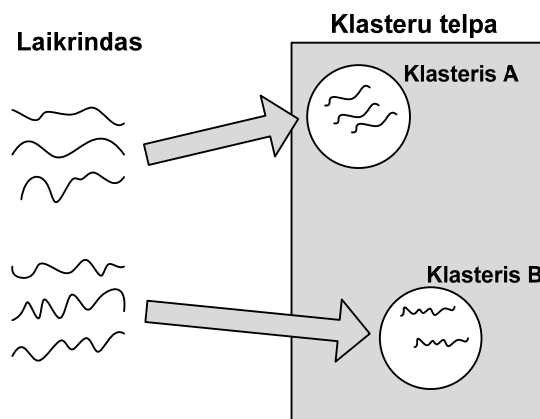
- (1) inicializē klasteru centrus w_j (j - nepieciešamo klasteru skaits uzdevuma risināšanai);
- (2) grupē visus apmācības izlases punktus ap tuvākā klastera centru, t.i., katru punktu x_i saista ar klasteru j^* , kuram

$$\|x_i - w_{j^*}\| = \min_j \|x_i - w_j\|;$$
- (3) izskaitļo jaunus klasteru centrus, t.i., visiem w_j izskaitļo :

$$w_j = \frac{1}{m_j} \sum_{x_i \in \text{klasterim } j} x_i,$$
 kur m_j – klasterim j piederošo punktu skaits;
- (4) atkārti (2) soli tik ilgi, kamēr iterāciju laikā nemainās klasteru centru vērtības.

Algoritma darbības rezultātā tiek noteikti galīgie klasteru centri w_j , ievērojot nosacījumu, ka attālumu kvadrātu summai starp visiem punktiem, kas pieder grupai j , un klastera centru ir jābūt minimālai. Laikrindu klasterizācijas būtība parādīta 3. attēlā.

Otrajā eksperimenta etapā laikrindu analizē tika izmantots k-vidējais klasterizācijas algoritms diviem klasteriem (bankrotu datiem un nebankrotu datiem). Algoritma darbības rezultātā vienam klasterim tika attiecināti laikrindu Uzņēmums1, Uzņēmums2 un Uzņēmums3 dati, bet otram klasterim - Uzņēmums47, Uzņēmums48 un Uzņēmums49 dati. Iegūtie klasteru centri ir šādi: [9; 5; 209; 63; 31] un [4; 4; 103; 95; 3].



3. attēls. Laikrindu klasterizācija ar k-vidējais klasterizācijas algoritmu
Avots: adaptēts no (7.)

Tādējādi, eksperimentu rezultātā tika izdarīts secinājums, ka konkrēto bankrota datu laikrindu klasterizācijas rezultāti ar k-vidējais algoritma palīdzību atbilst iegūtajiem rezultātiem ar LCSS metodi. Tas deva pārliecību, ka bankrotu datu kā laikrindu klasterizācijas rezultāti ir adekvāti LCSS.

Secinājumi un priekšlikumi

Darbā tika salīdzinātas vairākas metodes bankrotu datu analīzes nolūkā un izvērtētas to iespējas. Agrīnā jeb empīriskā pieeja izmanto atsevišķus finanšu rādītājus, pēc kuriem var spriest par potenciāla bankrota draudiem. Pētījumi pārsvarā tika veikti ar klasiskās statistikas metožu palīdzību - diskriminantanalīzes izmantošanu. Arī mākslīgo neironu tīklu izmantošanas potenciāls datu analīzes veikšanā ir nenovērtējams. Jebkurš neironu tīklu modelis tiek izmantots kā patstāvīga zināšanu uzdošanas sistēma, ar kuras palīdzību var realizēt noteiktas datu analīzes un apstrādes funkcijas. Iepriekšējos pētījumos tika secināts, ka konkrētai bankrotu datu izlasei neironu tīklu metode efektīvāk veic bankrotu diagnostiku nekā tradicionālā diskriminantanalīze.

Vispārīgā gadījumā laikrinda tiek traktēta kā datu virkne noteiktā laika intervālā. Laikrindu analīze ir pietiekami labi pazīstams uzdevums, taču pēdējā laikā tiek veikti pētījumi ar nolūku mēģināt pielietot klasterizāciju laikrindu analīzē - laikrindu datu sadalīšanu līdzīgās grupās. Laikrindu attēlošanas klasteru formā galvenā motivācija ir laikrindu raksturlielumu labāka izpratne. Darba galvenais mērķis bija izpētīt laikrindu klasterizācijas izmantošanas metodoloģiju, apzināt laikrindu līdzības novērtējumus un izmantot tos laikrindu klasterizācijas rezultātu novērtēšanā. Par līdzības novērtējumu tradicionāli tiek izmantota Eiklīda distance, taču pēdējā laikā līdzības mēra noteikšanai bieži izmanto vairāk komplicētas metodes, piemēram, garākās kopīgās virknes metodi.

Dotajā darbā tika veikti divi uzdevumi. Pirmais bija noteikt laicrindu līdzības novērtējuma iespējas. Eksperimenta mērķis bija pārbaudīt LCSS metodes piemērotību bankrotu datu kā laicrindu līdzības novērtēšanai. Otrais uzdevums bija izpētīt klasterizācijas algoritmu iespējas laicrindu klasterizācijā. Tika izmantots klasiskais klasterizācijas algoritms - k-vidējais un eksperimentu rezultātā izdarīts secinājums, ka konkrēto bankrotu datu kā laicrindu klasterizācijas rezultāti ar k-vidējais algoritma palīdzību atbilst iegūtajiem rezultātiem ar LCSS metodi. Tas dod pārliecību, ka laicrindu klasterizācijas rezultāti ir adekvāti un potenciāli iespējams bankrotu datu analizē izmantot laicrindu klasterizācijas iespējas. Tomēr pilnīgai izpratnei ir nepieciešams veikt padziļinātu laicrindu izmantošanas iespēju izpēti bankrotu datu analizē.

Izmantotā literatūra un avoti

1. Agrawal R., Faloutsos C., Swami A. Efficient similarity search in sequence databases. Proc. 4th Int. Conf. On Foundations of Data Organizations and Algorithms, 1993. – Chicago. pp. 69-84.
2. Altman E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance, vol. 13, p.589-609.
3. Atiya A. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. IEEE Transactions on Neural Networks, Vol. 12, No. 4, p. 929-935.
4. Back B., Laitinen T., Sere K. (1996). Neural networks and bankruptcy prediction: funds flow, accrual ratios and accounting data. Advances in Accounting 14, p. 23-37.
5. Beaver R. (1966). Financial ratios as predictors of failure, Empirical Research in Accounting: Selected Studies, J.Accounting Research, vol. 4, p.71-111.
6. Grabusts P. (2006). Bankrotu datu analīzes metožu iespēju izpēte. Starptautiskās zinātniskās konferences materiāli “Tautsaimniecības attīstības iespējas un problēmas”, Rēzekne, 24. marts, lpp. 55-63.
7. Keogh E., Lin J., Truppel W. Clustering of time series subsequences in meaningless implications for previous and future research. Proc. of the 3rd IEEE International Conference on Data Mining, 2003. – pp. 115 – 122.
8. Kirchgassner G., Wolter J. Introduction to modern time series analysis. – Berlin:Springer, 2007, 274 p.
9. Odom M., Sharda R. (1990). A neural network model for bankruptcy prediction. In Proc. Int. Joint Conf. Neural Networks, San Diego, CA.
10. Rudorfer G. (1995). Early bankruptcy detecting using neural networks. APL Quote Quad, ACM New York, vol. 25, N. 4, p. 171-178.
11. Tam K., Kiang M. (1992). Managerial applications of the neural networks: The case of bank failure predictions. Management Science, vol. 38, p.416-430.
12. Tsay R.S. Analysis of financial time series. – John Wiley & Sons, 2002, 448 p.
13. Vlachos M., Gunopulos D. Indexing time series under condition of noise. Data mining in time series database: Series in machine perception and artificial intelligence. – World Scientific Publishing, 2004. Vol.57, pp. 67-100.
14. <http://godefroy.sdf-eu.org/apl95/ratios95.zip> - skatīts 26.02.2010.

Summary

Under the conditions of the modern free market early diagnostics of unfavourable development trends of company's activity or bankruptcy becomes a matter of great importance. However, there is no general method which would allow one to forecast unfavourable consequences with a high degree of confidence. This paper focuses on the analysis of the approaches that can be used to perform an early bankruptcy diagnostics: multivariate discriminant analysis (MDA) and neural network based approach. The neural network based approach is analysed using a multilayer feed forward network with back propagation learning algorithm. Experiments carried out validate the use of such networks in the given class of tasks.

A time series is a sequence of real data, representing the measurements of a real variable at time intervals. Time series analysis is a sufficiently well-known task, however, recently research activities have been carried out with the purpose to try to use clustering for the intentions of time series analysis. The main motivation for representing time series in a cluster form is to represent the main characteristics of the data in a better way.

Time series clustering approach has become popular and its feasibility for bankruptcy data analysis is being investigated. Experiments carried out validate the use of such methods in the given class of tasks. As a novelty, an attempt to apply time series clustering method to the analysis of bankruptcy data is made.

In the present research paper two tasks have been completed. The first task was to define time series similarity measures. It has been established that LCSS method gives better results in the detection of bankruptcy data time series similarity than the Euclidean distance. The second task was to explore the facilities of the classical k-means clustering algorithm in time series clustering. As a result of the experiment a conclusion has been drawn that the results of time series clustering using k-means algorithm correspond to the results obtained with LCSS method, thus the clustering results of the specific bankruptcy data time series are adequate.