

VIRÁG Miklós – KRISTÓF Tamás

TÖBBDIMENZIÓS SKÁLÁZÁS A CSŐDMODELLEZÉS BEN

A tanulmány a csődmodellezési szakirodalomban alulreprezentált, ugyanakkor kiváló klaszterezési teljesítményt nyújtó többdimenziós skálázás csődmodellezési alkalmazását mutatja be empirikus vizsgálattal alátámasztva. Az empirikus vizsgálat alapján levonható a következtetés, hogy a többdimenziós skálázás a logisztikus regresszióelemzéssel kombinálva klasszifikációs képességben és interpretálhatóságban bármelyik alternatív csőd-előrejelzési módszerrel felveszi a versenyt.

Kulcsszavak: csődmodellezés, csődelőrejelzés, többdimenziós skálázás

A Bazel II Tőkeegyezmény bevezetésével a csődelőrejelzésre felhasználható sokváltozós statisztikai eljárások alkalmazásának jelentős erősödése figyelhető meg, elsősorban a pénzintézetek tevékenységi körében. A sokváltozós statisztikai eljárásokon alapuló csődelőrejelzésre általában a diszkriminanciaanalízist, a logisztikus regresszióelemzést, a rekurzív particionáló algoritmust (döntési fa) és/vagy a neurális hálókat alkalmazzák. A *Vezetéstudomány* olvasói a négy eljárás leírásával és azok előrejelző erejének összehasonlító empirikus vizsgálatával egy korábbi folyóiratcikkben (Virág – Kristóf, 2006) már találkozhattak.

Jelen tanulmány a fenti négy eljárás mellé kínál egy újabb alternatívát: a többdimenziós skálázást. A többdimenziós skálázás csődelőrejelzési alkalmazásának még a nemzetközi szakirodalma is rendkívül szűk, véleményünk szerint azonban érdemes figyelmet szentelni neki, mivel klaszterezőképességben és interpretálhatóságban bármelyik korábbi módszerrel felveszi a versenyt, sőt: előzetes feltevéseink szerint megbízhatóbb csődmodellezést tesz lehetővé, mint bármelyik korábban említett eljárás. Egyetértünk vele, és az empirikus vizsgálat remélhetőleg az olvasót is meggyőzi Neophytou – Mar Molinero (2004) véleményéről, hogy a többdimenziós skálázás paradigmaváltást idézhet elő a csődelőrejelzés területén.

A folyóiratcikk tömören ismerteti a többdimenziós skálázás módszertanát, külön kitérve az általunk alkalmazott PROXimity SCALing (PROXSCAL) eljárásra. A többdimenziós skálázás csődelőrejelzési alkalmazását empirikus vizsgálat keretében,

a pénzügyi mutatókon alapuló vállalati csődelőrejelzés példáján mutatjuk be. A tanulmány rávilágít arra, hogy a skálázás adattömörítésre és vizualizációra egyaránt kiválóan alkalmas. A csődelőrejelzés által igényelt csődvalószínűsége meghatározása és a csődmodell végfelhasználói számára egzakt döntési szabály kidolgozása egy kiegészítő eljárással: a logisztikus regresszióelemzés segítségével történik. A logisztikus regresszióval kombinált skálázómodell csődelőrejelzési teljesítményét a szokásos besorolási pontosság mutatókkal, a ROC-görbével és a ROC-görbe alatti területtel értékeljük. A kidolgozott csődmodell magas szintű megbízhatósága az előrejelzés-készítés területén hosszú idő óta érvényes alapvetést is igazolja, ami alapján a módszerkombinációk megfelelő alkalmazása megbízhatóbb előrejelzést tesz lehetővé, mint egy-egy módszer kizárólagos alkalmazása (Gáspár – Nováky, 2002).

A többdimenziós skálázás rövid bemutatása

A többdimenziós skálázás olyan statisztikai eljárás, amely az adatok közötti különbségeket vizualizálja, és az adatok rejtett struktúráját vizsgálja. Az adatok közötti rejtett összefüggések kimutatása a nem szakértők számára is közérthető modellezést tesz lehetővé. A skálázás az eredeti adatok között mért különbségből nyer információt, és származtat koordinátákat egy skálatérképen.

Az eljárás kiinduló feltételezése, hogy létezik a megfigyeléseknek egy kvantitatív reprezentációja.

A skálázómodellekben az objektumok az állapotter pontjaiként jelennek meg olyan módon, hogy a hasonló objektumok közel kerülnek egymáshoz (Füstös et al., 2004). Ennek megfelelően a többdimenziós skálázás vizuálisan különíti el a fizetőképes és a fizetésektelen vállalatokat. A skálázás feladata, hogy a minimális dimenziószámú térben olyan ponthalmazt találjon, hogy a térbeli távolságok monoton függvényei legyenek az adatok közötti különbségeknek.

A skálázás a megfigyelési egységek értékei alapján számított különbségeket kifejező adatok mátrixából indul ki. Az eljárás a hasonlóságok és különbségek alapján helyezi el a pontokat a megfelelő térpozícióba. Az algoritmus megtartja az ordinális kapcsolatokat, vagyis a nagy különbséggel rendelkező pontpárok az ábrázolt térben is messze kerülnek egymástól, míg az alacsonyabb különbséggel rendelkező pontpárok a térben közel helyezkednek el egymáshoz (Kruskal – Wish, 1978). Az eljárás a származtatott koordináták közötti távolságokat összeveti az eredetileg ismert különbségekkel, és törekszik az eltérés minimalizálására. A különbségeket általában euklideszi távolsággal méri.

A többdimenziós skálázás a hagyományos statisztikai eljárásokkal ellentétben nem támaszt az adatokkal szemben követelményeket. Annyi az elvárás, hogy az adatok üzenetet hordozzanak magukban, és azok azonos mértékegységűek legyenek.

Ordinális skálán kvantifikált változók esetén nem-metrikus skálázásról, intervallum vagy arányskálán kvantifikált változók esetén metrikus skálázásról beszélünk (Kovács, 2006). Mivel a csődelőrejelzésre felhasznált pénzügyi mutatók mindegyike arányskálán kvantifikált változó, ezért esetünkben a metrikus skálázás esete áll fenn. Ez a tulajdonság a skálázás szempontjából kedvező, mivel a nem-metrikus skálázás gyakran eredményez lokális minimumot vagy nem konvergál.

A többdimenziós skálázás térkép formájú statisztikai reprezentációt állít elő. A térkép dimenziószámának megválasztása rendkívül fontos bármilyen elemzés végrehajtása előtt, a szakirodalomban azonban nincsen jól bevált recept arra vonatkozóan, miként választható ki az optimális dimenziószám. Ebből következően többféle megközelítést célszerű figyelembe venni.

A többdimenziós skálázás előnye, hogy eredményeképpen könnyen interpretálható statisztikai térkép adódik. Az eredmények a nélkül is értékelhetők, hogy a felhasználó mélységében tisztában lenne az eljárás statisztikai hátterével. Az *outlier* megfigyelések jelenléte és az alapadatok közötti korreláció a skálázás számára nem jelent problémát. Nincsenek eloszlási feltételek sem.

A többdimenziós skálázás rendkívül gazdag eljárás-családdal rendelkezik. Az általunk alkalmazott eljárás kiválasztásánál az képviselte a döntő szempontot, hogy egyrészt találjunk a nemzetközi szakirodalomban legalább egy példát annak csődmodellezési alkalmazására, másrészt rendelkezünk szoftverrel az eljárás lefuttatására. Ezek alapján esett választásunk a PROXSCAL-eljárásra, amelyet korábban Neophytou – Mar Molinero (2005) alkalmazott csődmodellezésre.

A PROXSCAL-eljárás

A PROXSCAL-eljárás a különbözőségi adatokon hajt végre többdimenziós skálázást oly módon, hogy az objektumok legkisebb négyzetes reprezentációja az eredetinel lényegesen alacsonyabb dimenziójú térben valósuljon meg. Az eljárás matematikai hátteréről *Commandeur – Heiser* (1993) ad részletes leírást.

A PROXSCAL-eljárás a transzformált különbségek és az objektumok közötti távolságok súlyozott négyzetes hibáját minimalizálja az alábbi veszteségfüggvény minimalizálásával:

$$\sigma^2 = \frac{1}{m} \sum_{k=1}^m \sum_{i < j}^n w_{ijk} \left[\hat{d}_{ijk} - d_{ij}(X_k) \right]^2$$

ahol: m = források száma,

\hat{d}_{ijk} = objektumok száma,

= transzformált különbségek,

$d_{ij}(X_k)$ = objektumok közötti euklideszi távolság az egyedi térkoordinátákat tartalmazó

X_k = soraiban lévő koordinátákra vonatkozóan,

i, j, k = futóindexek.

A PROXSCAL-algoritmus nem tér el jelentősen a skálázóeljárásoknál megszokott módszertől. Első lépés X_k kezdeti konfigurációjának megtalálása. A kezdeti konfiguráció kialakítható a szimplex, a *Torgerson*-eljárásokkal, véletlen módon vagy a felhasználó által meghatározott értékekkel. Ezt követi a kezdeti konfiguráció módosítása a közös tér szempontjából optimális konfiguráció megtalálása érdekében. Ezután a módosított X_k konfiguráció alapján megtörténik a transzformált különbségek optimalizálása. A veszteségfüggvény értékelése minden lépésben végrehajtásra kerül. A konfiguráció és a transzformált különbségek addig módosulnak, ameddig sikerül a minimumot megtalálni. A metrikus skálázás általában globális minimumot ad.

A PROXSCAL-eljárással elkészített modell illeszkedési jóságát a *normalizált raw stressz* mutató segítségével mérhetjük. Ahogyan az más skálázóeljárásoknál is megszokott, a minél alacsonyabb stresszmutató jelzi a modell jobb illeszkedését. A különböző dimenziószám esetén kiszámított stresszértékek összehasonlításával a dimenziók „hozzáadott értékéről” szerezhetünk információt.

A minta összetétele, magyarázó változók, adatelőkészítés

A rendelkezésre álló SPSS statisztikai programcsomag két jelentős korlátot emelt a skálázás végrehajthatósága tekintetében. Egyfelől az SPSS-ben futtatható ALSCAL- és PROXSCAL-eljárások egyaránt maximum 100×100-as különbözőségi mátrixszal képesek dolgozni, másfelől a dimenziók száma maximum hat lehet. A csődmodellezésben tapasztalataink alapján a hat dimenzió nem probléma, a különbözőségi mátrix korlátja azonban azt jelenti, hogy maximum száz megfigyelés vagy száz változó között lehet különbözőségeket számítani.

Amennyiben azt tűztük volna ki célul, hogy különböző pénzügyi mutatók hatását és egymástól vett eltérését vizsgáljuk a fizetőképesség szempontjából, akkor a későbbiekben ismertetett 31 pénzügyi mutatóhoz tetszőleges számú vállalat lenne kapcsolható. Jelen folyóiratcikk azonban arra kíváncsi, hogy miként különíthetők el legjobban a fizetőképes és a fizetéseképtelen vállalatok egymástól, ezáltal a többdimenziós koordinátákat magukhoz a vállalatokhoz kell skálázni. Ez azonban csupán száz vállalat megfigyelését teszi lehetővé, ami kis minta ugyan a megalapozott következ-

etések levonásához, elegendő azonban a többdimenziós skálázás működésének megértéséhez és az előnyök felismeréséhez.

A fenti korlátokat figyelembe véve száz vállalat 2004. évi mérlegei és eredménykimutatásai kerültek összegyűjtésre. A százalékos mintából 50 fizetőképes és 50 fizetéseképtelen vállalat volt. Mindegyik vállalat százmillió Ft feletti nettó árbevételrel és mérlegfőösszeggel rendelkezett. A fizetéseképtelenséget a 2005. évben megindított csődeljárás, felszámolási eljárás vagy végelszámolás jelentette. A 2004. évi mérlegek és eredménykimutatások a fizetéseképtelen társaságok esetén a fizetéseképtelenség bejelentéséhez viszonyított utolsó beszámolót jelentik, vagyis maximum 12 hónap lehet a különbség az éves beszámoló fordulónapja és a fizetéseképtelenség deklarálása között.

A száz megfigyelés 7 nemzetgazdasági ágazatból, azon belül 30 ágazatból (kétjegyű TEÁOR-kód) és azon belül 66 szakágazatból (négyjegyű TEÁOR-kód) származott. A 66-féle szakágazatból származó száz vállalat kezelése különös gondosságot igényelt a csődmodellezésben, hiszen köztudott, hogy a pénzügyi mutatókat közvetlenül nem szabad ágazatok között összehasonlítani. A modellezés ezért a sokasági szakágazati átlagoktól való eltéréseken került végrehajtásra.

A magyarázó változók a fizetőképességgel bizonyítottan összefüggésben lévő jövedelmezőségi, forgási sebesség, eladósodottsági, tőkeszerkezeti, likviditási, cash flow, méret és éves növekedési mutatókból kerültek ki (Virág, 1996). 31 pénzügyi mutató került kiválasztásra, szakmai szempontok alapján, amelyeket az 1. táblázat foglal össze.

1. táblázat

Az alkalmazott pénzügyi mutatók neve és számítási eljárása

Mutató neve	Számítási eljárás
Vagyonarányos nyereség (ROE)	Adózott eredmény / Átlagos saját tőke
Eszközhatékonysági mutató (ROA)	Adózott eredmény / Átlagos mérlegfőösszeg
Árbevétel-arányos nyereség (ROS)	Üzemi (üzleti) tevékenység eredménye / Értékesítés nettó árbevétele
Árbevétel-arányos EBITDA	(Üzemi [üzleti] tevékenység eredménye + Értékcsökkenési leírás) / Értékesítés nettó árbevétele
EBITDA-jövedelmezőség	(Üzemi (üzleti) tevékenység eredménye + Értékcsökkenési leírás) / Átlagos mérlegfőösszeg
Átlagos eszközállomány forgása	Értékesítés nettó árbevétele / (Átlagos mérlegfőösszeg / 365)
Készletek forgási sebessége	Értékesítés nettó árbevétele / (Átlagos készletállomány / 365)
Vevők forgási sebessége	Értékesítés nettó árbevétele / (Átlagos vevőállomány / 365)
Saját vagyon aránya	Saját tőke / Mérlegfőösszeg
Hosszú távú eladósodottság	Hosszú lejáratú kötelezettségek / (Saját tőke + Hosszú lejáratú kötelezettségek)
Befektetett eszközök saját finanszírozása	Saját tőke / Befektetett eszközök

Mutató neve	Számítási eljárás
Eladósodottság mértéke	Kötelezettségek / Mérlegfőösszeg
Bonitás	Kötelezettségek / Saját tőke
Befektetett eszközök idegen finanszírozása	Hosszú lejáratú kötelezettségek / Befektetett eszközök
Tőkeellátottsági mutató	(Befektetett eszközök + Készletek) / Saját tőke
Forgóeszköz-arány	Forgóeszközök / Mérlegfőösszeg
Pénzeszközök aránya	(Pénzeszközök + Értékpapírok) / Forgóeszközök
Nettó forgótőke-arány	(Forgóeszközök – Rövid lejáratú kötelezettségek) / Mérlegfőösszeg
Likviditási ráta	Forgóeszközök / Rövid lejáratú kötelezettségek
Likviditási gyorsráta	(Forgóeszközök – Készletek) / Rövid lejáratú kötelezettségek
Készpénzlikviditás	(Pénzeszközök + Értékpapírok) / Rövid lejáratú kötelezettségek
Dinamikus likviditás	Üzemi (üzleti) tevékenység eredménye / Rövid lejáratú kötelezettségek
Vevők / Szállítók aránya	Vevőkövetelések / Szállítói kötelezettségek
Dinamikus jövedelmezőségi ráta (bruttó)	(Adózott eredmény + Értékcsökkenési leírás) / Átlagos mérlegfőösszeg
Cash flow / összes tartozás	(Adózott eredmény + Értékcsökkenési leírás) / (Hosszú lejáratú kötelezettségek + Rövid lejáratú kötelezettségek)
Cash-flow / nettó árbevétel	(Adózott eredmény + Értékcsökkenési leírás) / Értékesítés nettó árbevétele
Mérlegfőösszeg nagysága	log (Mérlegfőösszeg)
Éves árbevétel nagysága	log (Értékesítés nettó árbevétele)
Árbevétel növekedési üteme	Értékesítés nettó árbevétele tárgyidőszak / Értékesítés nettó árbevétele előző időszak
Üzemi (üzleti) eredmény növekedése	Üzemi (üzleti) tevékenység eredménye tárgyidőszak / Üzemi (üzleti) tevékenység eredménye előző időszak
Adózott eredmény növekedése	Adózott eredmény tárgyidőszak / Adózott eredmény előző időszak

Az adatgyűjtést követte az adatok feldolgozása és azok modellezésre történő előkészítése. Ez sokszor nehezebb feladat, mint maga a modellezés, hiszen itt jelentkeznek a megfigyelésekkel és/vagy a változókkal kapcsolatban előzetesen nem várt problémák.

Három pénzügyi mutató (saját vagyonarányos nyereség »ROE«, üzemi [üzleti] eredmény növekedése, adózott eredmény növekedése) számítása során kellett kettős negatív osztásokat kezelni. A ROE 20 vállalatnál, az üzemi (üzleti) eredmény növekedése 26 vállalatnál, az adózott eredmény növekedése 25 vállalatnál származott volna negatív számlálóból és negatív nevezőből.

A ROE esetén ez azt jelenti, hogy a mintában olyan vállalatok is szerepelnek, amelyeknek a tartozásai egyrészt meghaladják a vagyont, másrészt veszteséggel zárták az évet, és a mutató ezeket elferdítve pozitív jövedelmezőséget mutat.¹ Természetesen ez túlnyomórészt a fizetéseképtelen megfigyelésekre jellemző. A kétféle eredménykategória növekedése pedig mindkét évet negatív eredménnyel záró (akár a tárgyévben tovább romló üzemi vagy adózott eredményű) vállalatok esetében eredményez pozitív növekedést. A gyakorlatban jól bevált adatbányászati technika, hogy ilyenkor a kettős negatív értékekkel rendelkező megfigyelések mutató-

számértékét a többi megfigyelés adott mutatójának minimumával helyettesítik, de tekintettel a kis mintára és a viszonylag nagy számú érintett vállalatra, ez a három mutató inkább kikerült az empirikus vizsgálatból.

Az adatelőkészítés feladatkörébe tartozott az egyedi pénzügyi mutatók négyjegyű TEÁOR-kódokra számított sokasági szakágazati átlagokkal történő korrekciója. A korrekció az alábbi képlet segítségével történt:

$$\frac{\text{Egyedi mutatószámérték} - \text{Szakágazati átlag értéke}}{\text{Szakágazati átlag értéke}}$$

A szakágazati átlagokkal való korrekció megteremtette az összemérhetőséget egymástól jelentősen eltérő tevékenységi körökkel rendelkező vállalatok között. Ettől a ponttól kezdve a vizsgálat nem a mutatószámértékek nagyságára, hanem azoknak a saját szakágazatukra jellemző átlagokhoz viszonyított eltérésére vonatkozik. Ezáltal a modellek időbeni stabilitása is javul, hiszen az átlagokhoz képest jobb vagy rosszabb teljesítmény évek múlva is releváns szempontnak bizonyulhat a fizetőképesség megítélése során.

A csődmodellezés során kulcsfontosságú kérdés az outlier megfigyelések csonkolása. Tekintettel azonban

arra, hogy a skálázás érzéketlen az *outlierekre*, nem végeztünk csonkolást. Néhány megfigyelés esetén az ábrákon látszani fog, hogy a kiugró értékek a koordinátákban a többséghez képest viszonylag távol helyezkednek el.

A többdimenziós skálázás alkalmazási feltételei közé tartozik az adatok standardizálása. Ez megfigyelésként és változónként egyaránt lehetséges. Jelen vizsgálatban a standardizálás változónként történt egységnyi szórásra.

A skálázómodell bemutatása

A többdimenziós skálázás a skálatérkép koordinátáit kizárólag a pénzügyi mutatók közötti különbözőségeken alapján számítja, az tehát nem veszi figyelembe, hogy a megfigyelések fizetőképesek vagy fizetéseképtelenek. A különbözőségeket arányskálán, metrikus skálázással, euklideszi távolság számításával határoztuk meg az egységnyi szórásra standardizált alapadatokon. Az eljárástól 2-6 dimenziós megoldásokat kértünk.

A kezdeti konfiguráció a szimplex módszerrel került kialakításra. A minimum stresszértéket 0,0001-nek, a maximális iterációk számát 500-nak állítottuk be. A stresszkonvergenciát 0,0001-nek választottuk meg, ami azt jelenti, hogy az iteráció akkor áll le, amikor a *normalizált raw stressz* értéke a megadott konvergencia szintnél kevésbé javul. A hatdimenziós megoldás esetén 34 iteráció után az algoritmus a konvergencia követelménynek megfelelően leállt, ekkor a *normalizált raw stressz* értéke 0,00759 volt. Ez a rendkívül alacsony érték jó modellilleszkedést mutat. Az illeszkedési jóságról a *Tucker*-féle kongruencia-együttható segítségével is meggyőződhetünk. Minél közelebb van a kongruencia-együttható 1-hez annál jobb az illeszkedés. A hatdimenziós skálázó modell kongruencia-együtthatója 0,99620 volt, ami szintén a magas szintű illeszkedést mutatja.

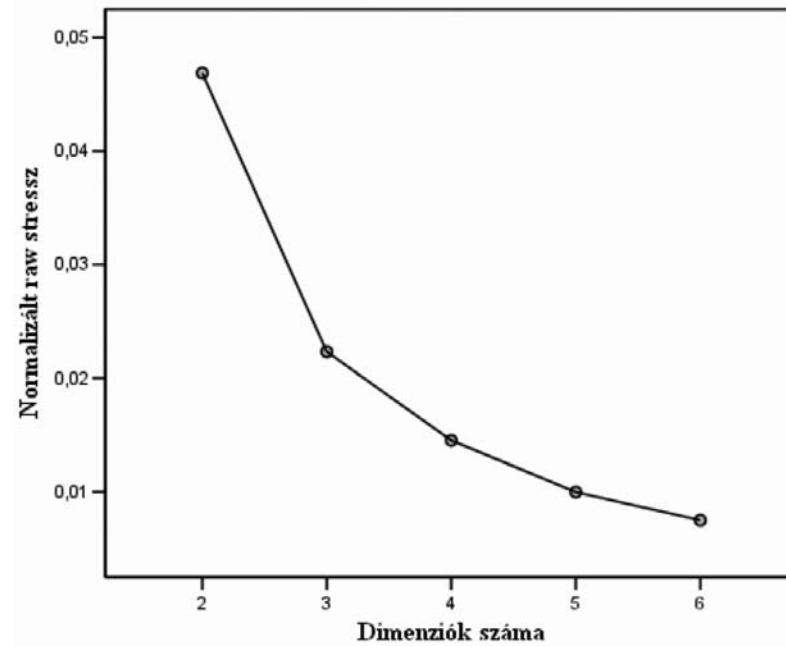
A többdimenziós skálázás során kulcsfontosságú probléma a dimenziószám optimális megválasztása. Mivel erre a szakirodalomban nem található egyértelmű iránymutatás, ezért többféle szempontból közelítettük meg a dimenziószám megfelelő megválasztását. Az SPSS-ben a PROXSCAL-eljárás 2 és 6 dimenzió közötti megoldásokat képes előállítani.

Első megközelítésben az illeszkedési jóság mutató és a dimenzionalitás kapcsolatát vizsgáltuk. Az 1. ábrán felrajzoltuk a *normalizált raw stressz mutató* értékét 2-6 dimenziószám esetén. Az ábrát ugyanúgy kell

értelmezni, mint amikor a főkomponens-elemzés esetén ábrázoljuk a saját értékeket a komponensek számának függvényében, és azt a komponensszámot tekintjük elégségesnek, amikor a saját értékek csökkenése egyértelműen lassul.² Az 1. ábra alapján már a harmadik dimenzió is feleslegesnek tűnik.³

1. ábra

A normalizált raw stressz értékek különböző dimenziószámok esetén

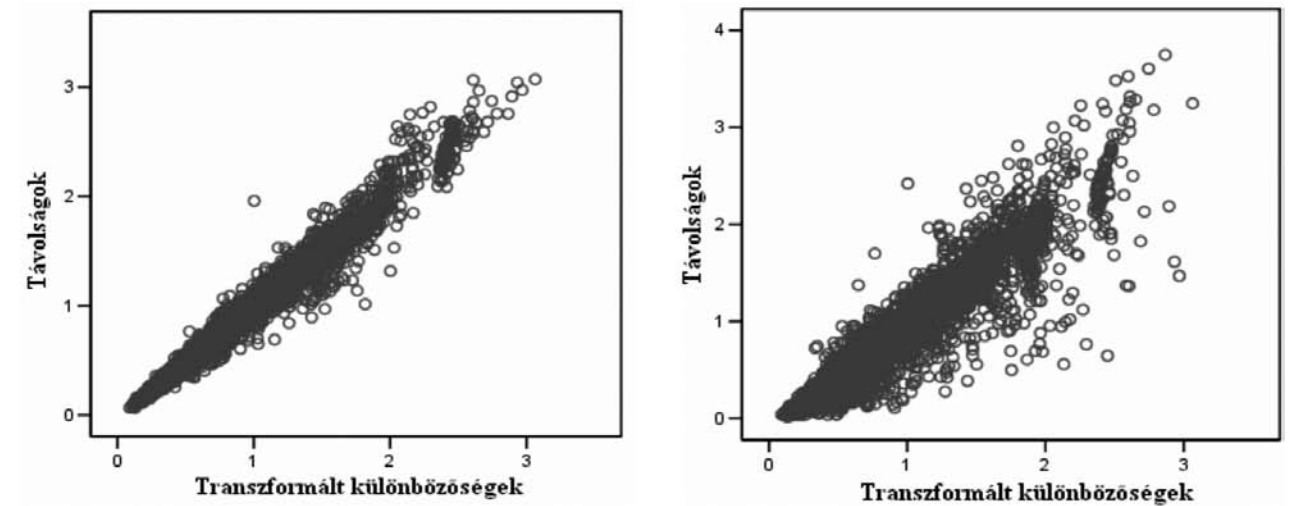


A dimenziószám megválasztásához további felhasználható elemzési szempont (ami egyben a skálázómodell által kihozott megoldás minőségének megítéléséhez is jól alkalmazható) az eredeti adatok és a skálázómodell eltéréseinek vizsgálata. Erre a gyakorlatban jól bevált elemző eszköz az ún. *Shepard*-diagram (*Groenen-van de Velden*, 2004) tanulmányozása, amely a transzformált különbözőségeket és a távolságok megfeleltetését ábrázolja a reziduumban segítségével. A *Shepard*-diagram segítségével kimutathatók az *outlier* megfigyelések és a skálázómodell eredményezte hibák.

A *Shepard*-diagramot minden dimenzióra vonatkozóan célszerű számítógéppel előállítani. Minél inkább összetartanak a pontok, annál jobb a választott dimenziószám. A 2. ábra bal oldali diagramja a hatdimenziós megoldás reziduumaikat, a jobb oldali diagramja a kétdimenziós megoldását mutatja be. Az ábra egyben igazolja a skálázómodell magas szintű teljesítményét. A dimenziószám csökkentésével a transzformált különbözőségeket és a távolságok ábrája lényegesen szerkezetesebb.

2. ábra

Reziduumban ábrázolása a hat- és kétdimenziós megoldás esetén (Shepard-diagram)



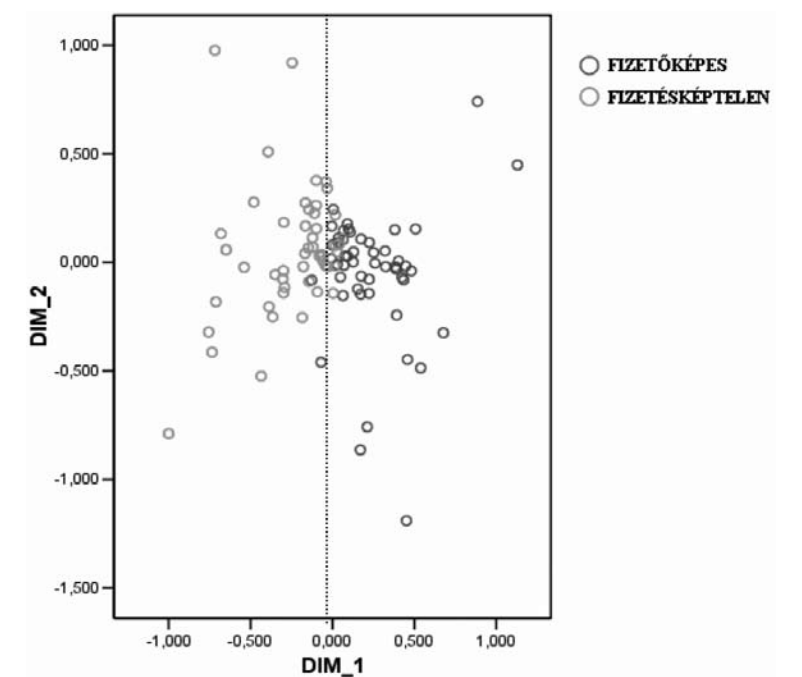
A reziduumban vizsgálata alapján tehát mind a hat dimenzióra szükség van. Mivel az objektív statisztikai elemzési szempontok alapján nem lehet egyértelműen eldönteni a dimenziószámot, ezért a szakmai szempontok alapján történő interpretációt vettük a továbbiakban irányadónak. Valószínűsítettük, hogy a csődelőjelzéshez nincs szükség mind a hat dimenzióra, de jelentős információvesztést sem akartunk. A választott módszer a különböző dimenziókban létrejött koordináták dimenziópáronkénti ábrázolása, és azokból szabad szemmel összefüggések és relevancia megállapítása volt.⁴

A hat dimenzióhoz tartozó koordinátákat a PROXSCAL-eljárás biztosítja minden megfigyeléshez. Ezt követi a dimenziók értékelése csődelőjelzési szakmai szempontokból. Cél a fizetőképesek és a fizetéseképtelen megfigyelések minél jobb elkülönítését jelző dimenziókat megtalálni. A szakmai elemzést a későbbiekben logisztikus regressziószámítás követi. Ennek megfelelően három pontdiagram készült: az 1. és a 2., a 3. és a 4., illetve az 5. és a 6. dimenziópárok alapján. Lehetőség lett volna minden dimenziókombinációra pontdiagramot felrajzolni, az összefüggések megállapítására azonban esetünkben elegendőnek bizonyult, ha mindegyik dimenzió egyszer kerül ábrázolásra. A megfigyelésekhez kiszámított koordinátákat elláttuk az általunk ismert fizetőképesség tényével *dummy* változóként, hiszen ezt az információt a skálázás nem vette figyelembe. A diagramokon szürke körök jelölik a fizetőképesek és fekete körök jelölik a fizetéseképtelen vállalatok koordinátáit.

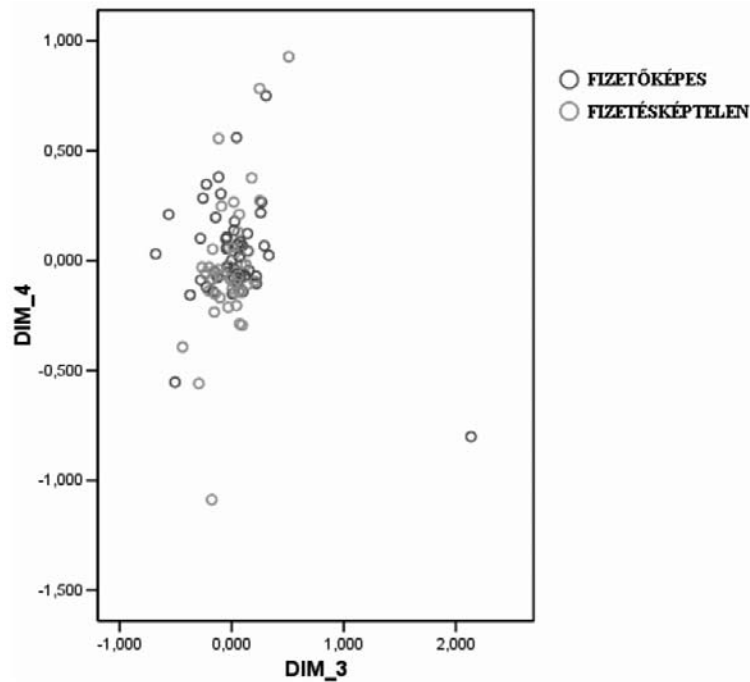
A hatdimenziós térkép 1. és 2. dimenzióra vonatkozó kivetítését tartalmazó 3. ábra nagyon fontos tulajdonságot jelez. Az tapasztalható, hogy az ábra bal oldalára összpontosulnak a fizetőképesek és a jobb oldalára a fizetéseképtelen megfigyelések. Szabad szemmel is látható, hogy az 1. dimenzió nagyon jól magyarázza a fizetőképesek és a fizetéseképtelen megfigyelések közötti különbséget, míg a 2. dimenzió gyakorlatilag semennyire.

3. ábra

A fizetőképesek és a fizetéseképtelen megfigyelések koordinátái az 1. és a 2. dimenzióban



A fizetőképes és a fizetésképtelen megfigyelések koordinátái a 3. és a 4. dimenzióban



4. ábra A 3. ábra alapján az a következtetés vonható le, hogy még az utóbbi években jelentősen megkritizált lineáris elválasztás is működőképesnek bizonyulhat a skálázással képzett koordináták alapján: egy az 1. dimenzió origójára állított függőleges egyenes ugyanis majdnem tökéletesen képes elválasztani a fizetőképes és a fizetésképtelen osztályokat.

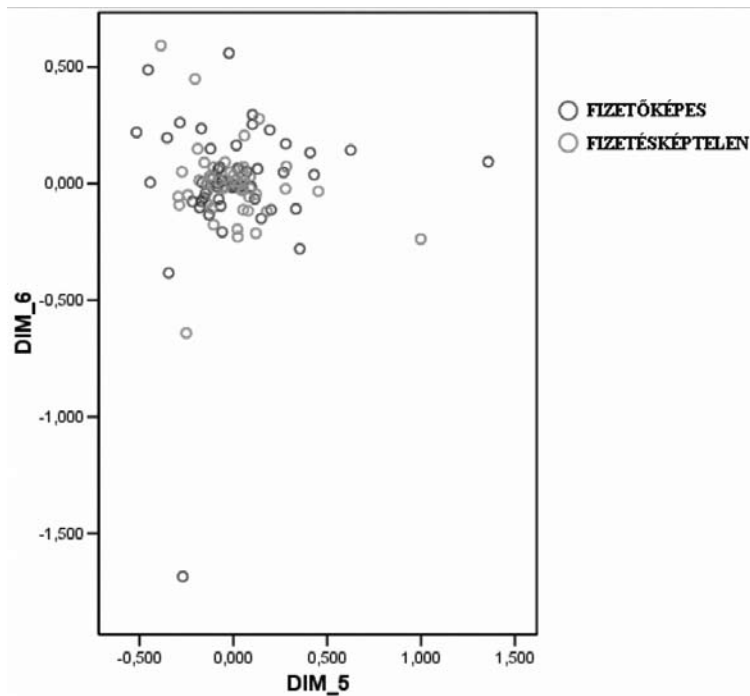
A 4. ábrán szereplő 3. és 4. dimenziók magyarázó ereje szabad szemmel már nem annyira nyilvánvaló, mint a korábbi dimenzióknál. Az ábra tanulmányozása alapján megállapíthatjuk, hogy talán a 4. dimenzió bírhat klasszifikációs képességgel, de annak kimutatása további vizsgálatokat igényel.

Az 5. és a 6. dimenziók közül szabad szemmel egyik sem tűnik szignifikánsnak. A fizetőképes és a fizetésképtelen megfigyelések meglehetősen összevissza szóródnak az 5. ábrán.

Szakmai elemzésünk alapján az sejthető, hogy az 1. dimenzió biztosan, a 4. dimenzió pedig talán jól magyarázza a várható fizetőképességet. A látványos ábrák azonban önmagukban csődelőrejelzésre még nem használhatók, hiszen a többdimenziós koordinátákból automatikusan nem származtathatók a hitelkockázat-kezelés szempontjából rendkívül fontos csődvalószínűségi értékek, nem olvasható ki a besorolási pontosság, és nem tudjuk, hogy az első dimenzió kívül melyik további dimenziók lehetnek szignifikánsak a fizetőképes és fizetésképtelen osztályok megkülönböztetése szempontjából. A probléma megoldására a logisztikus regresszióelemzést vettük segítségül.

5. ábra

A fizetőképes és a fizetésképtelen megfigyelések koordinátái a 5. és a 6. dimenzióban



Logisztikus regresszióelemzés a skálázó-modellen

A logisztikus regressziómodell a gyakran alkalmazott *forward stepwise* eljárással került kidolgozásra. A magyarázó változók a hat dimenzió koordinátái, minden megfigyelésre vonatkozóan, függő változó a fizetőképesség ténye. A fizetőképes megfigyeléseket 0-val, a fizetésképtelen megfigyeléseket 1-gyel jelöltük.

A *forward stepwise* eljárás egyesével lépteti be a szignifikánsnak talált magyarázó változókat a modellbe. Jelen vizsgálatban a Wald-féle bekerülési és kikerülési kritériumokat vettük figyelembe a megszokott 5% és 10% szignifikanciaszinteken. A konstanson kívül első lépésben az 1. dimenzió, második lépésben

A logisztikus regressziós modell együtthatói, változói és a változók tesztelése

Modellváltozó	β	Standard hiba	Wald-teszt	p-érték	exp(β) (95%-os CI)
DIM_1	37,130	11,316	10,767	0,001	1E+016 (3113947-6E+025)
DIM_4	11,651	4,719	6,095	0,014	114802,4 (11,040-1E+009)
DIM_6	4,836	2,413	4,016	0,045	125,941 (1,112-14260,249)
Konstans	-0,030	0,517	0,003	0,953	0,970

a 4. dimenzió, harmadik lépésben a 6. dimenzió került modellváltozóként felvételre. A *p*-értékekből látható, hogy a modellváltozók 95% valószínűségi szint mellett szignifikánsak (2. táblázat).

A logisztikus regressziós modellben szereplő változók szignifikanciája igazolta szubjektív vélekedésünket az 1. és a 4. dimenzióról. A 95% szignifikanciaszintbe „még éppen belefért” a 6. dimenzió is, azt korábban szabad szemmel nem találtuk relevánsnak. A modell egészének szignifikanciáját az *Omnibus*-féle χ^2 -próbaival teszteltük. Az empirikus χ^2 -érték 110,901 (szabadságfokok száma: 3), a *p*-érték 0,000. Ebből következően a logisztikus regressziós modell minden valószínűségi szint mellett szignifikáns.

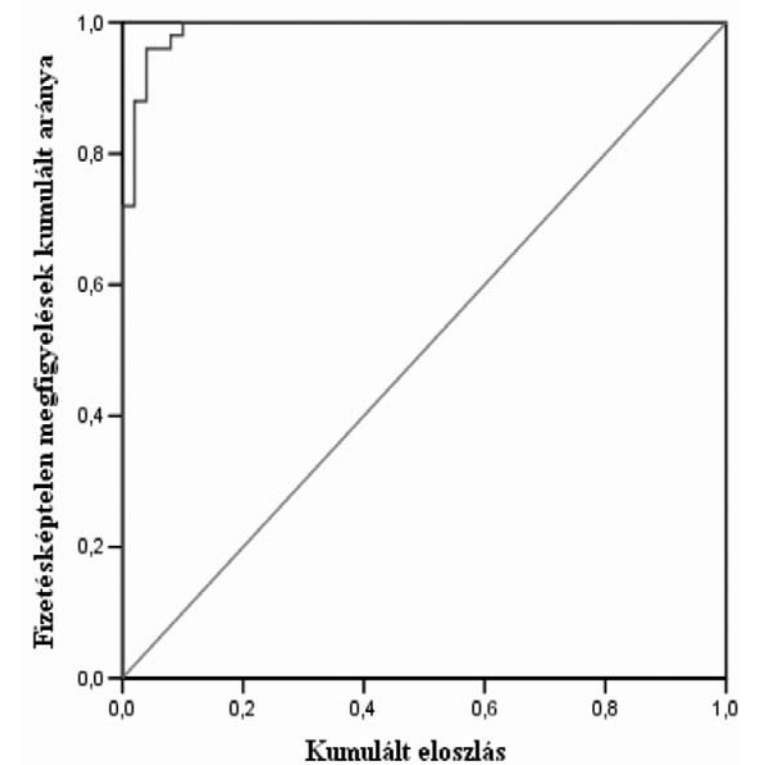
A *cut-off* érték a kiegyensúlyozott mintavételkor szokásos 50%-os szinten került megválasztásra. A besorolási pontosságokból látható, hogy a modell a fizetőképes és a fizetésképtelen megfigyeléseket egyaránt rendkívül magas szinten: 96, illetve 92% pontossággal képes helyesen besorolni. Ez ilyen kis mintán nagyon jó eredménynek számít. A logisztikus regressziós modellt elkészítettük kizárólag az 1. dimenzió modellváltozóval is, ekkor a háromváltozós modellnél tapasztalt 94% összbesorolási pontosság 89%-ra csökkent.

A csődelőrejelzés értékelése a klasszifikációs mátrixon kívül a ROC-görbével és a görbe alatti területtel is szükséges, hiszen az minden létező *cut-off* értéket figyelembe vesz. A ROC-görbe azt vizsgálja, hogy a modellek futtatásával kapott valószínűségi értékek mennyire jelzik megbízhatóan az output-kategóriába való tartozást, amennyiben az eredeti besorolás ismert. A vízszintes koordinátatengely a kumulált eloszlást, a függőleges koordinátatengely a fizetésképtelen megfigyelések kumulált arányát fejezi ki. A ROC-görbe referenciája a 45°-os egyenes, amely a véletlen találgatásnak felel meg. Annál jobb az értékelése valamely csődelőrejelzésnek,

minél jobban elválik a ROC-görbéje a 45°-os egyenestől. A 6. ábrán látható, hogy a csődelőrejelzés ROC-görbéje határozottan elválik a referenciavonaltól, és nagyon hamar tetőzik, vagyis a modellilleszkedés kitűnő.

6. ábra

A logisztikus regressziós modell ROC-görbéje



A ROC-görbéből számított objektív statisztikai mutató a görbe alatti terület nagysága.⁵ Amennyiben a görbe alatti terület 0,5 felett van, akkor az rendelkezik hozzáadott értékkel a véletlen találgatáshoz viszonyítva. Minél nagyobb valamely csődelőrejelzés ROC-görbe alatti területe, annál jobb. A ROC-görbe alatti terület 0,990, ami kiemelkedően magas szintű modellilleszkedést mutat.

Következtetések

A Bazel II Tőkeegyezmény bevezetésével a sokváltozós statisztikai alapokon nyugvó csőd-előrejelzési módszerek jelentős erősödése figyelhető meg. A folyóiratcikkéből látható, hogy a többdimenziós skálázásnak egyértelműen helye van a sokváltozós csőd-előrejelzési technikák között. A többdimenziós skálázás annak ellenére képes pontos klaszterezésre, hogy kizárólag a pénzügyi mutatókból származtatott különbözőségek mátrixából dolgozik, vagyis nem veszi figyelembe a fizetőképesség tényét a koordináták meghatározása során. Az empirikus vizsgálat alapján háromdimenziós megoldás született.

A kidolgozott csődmodell rendkívül jó teljesítménye alapján megállapíthatjuk, hogy igaza volt a *Neophytou – Mar Molinero* szerzőpárosnak, amikor azt állították, hogy a többdimenziós skálázás paradigmaváltást idézhet elő a csődelőrejelzés területén. Az empirikus vizsgálat alapján igazolást nyert, hogy a többdimenziós skálázás nagyon jó adattömörítő és vizuális klaszterező-eljárás. A többdimenziós koordinátákon lefutott logisztikus regresszióelemzéssel kiegészítve a többdimenziós skálázás a csődelőrejelzés minden követelményének megfelel. A 94% besorolási pontosság és a 99% ROC-görbe alatti terület még bennünket is meglepett a 100 elemű mintán.

A skálázómodell becslőképességét és az interpretálhatóságot némileg beárnyékolja a modell új megfigyeléseken történő alkalmazási nehézsége, hiszen a modell által nem ismert vállalatok pénzügyi mutatóiból automatikusan nem állíthatók elő hatdimenziós koordináták. A probléma úgy orvosolható, hogy az új vállalat adatait bele kell foglalni a mintába, és újrafuttatni a modellezést, azonban erre már csak a sokváltozós statisztikában jártas szakember képes, a késztermékre kíváncsi felhasználó általában nem.

Lábjegyzet

¹ Szemben olyan vállalatokkal, ahol csak az egyik tétel negatív. Olyan eset is előfordulhat, hogy a kettős negatív hatás kioltódása miatt kettős pozitív értékű vállalatnál is jobb jövedelmezőséget hoz ki eredményül a mutató.

² Ezt szokták könyök-szabálynak is nevezni.

³ Sok elemző a könnyű interpretáció végett eleve törekszik a kétdimenziós megoldásra.

⁴ Hasonló megoldást választott *Neophytou – Mar Molinero* (2004), akik az INDIVIDUAL SCALING (INDSCAL) eljárás által eredményezett koordinátákon alkalmaztak páronkénti grafikus ábrázolást.

⁵ A ROC-görbe alatti terület nagyságához hasonló információt biztosít a GINI-mutató, amely a 45°-os egyenes feletti területrészt arányosítja a tökéletes klasszifikációhoz.

Felhasznált irodalom

- Commandeur, J.J.F. – Heiser, W.J.* (1993): Mathematical derivations in the proximity scaling (PROXSCAL) of symmetric data matrices. Technical Report No. RR-93-03. Department of Data Theory, Leiden University, Leiden
- Füstös L. – Kovács E. – Meszéna Gy. – Simonné Mosolygó N.* (2004): Alakfelismerés. Sokváltozós statisztikai módszerek. Új Mandátum Kiadó, Budapest
- Gáspár, T. – Nováky, E.* (2002): Dilemmas for renewal of futures methodology, *Futures*, Vol. 34. No. 5. 365–379. old.
- Groenen, P.J.F. – van de Velden, M.* (2004): Multidimensional scaling. *Econometric Institute Report EI 2004-15*. Econometric Institute, Erasmus University of Rotterdam, Rotterdam
- Kovács E.* (2006): Pénzügyi adatok statisztikai elemzése. Budapesti Corvinus Egyetem, Pénzügyi és Számviteli Intézet, Budapest
- Kruskal, J.B. – Wish, M.* (1978): *Multidimensional Scaling*. Sage Publications, London
- Neophytou, E. – Mar Molinero, C.* (2004): Predicting Corporate Failure in the UK: A Multidimensional Scaling Approach. *Journal of Business Finance and Accounting*, Vol. 31. No. 5–6. 677–710. old.
- Neophytou, E. – Mar Molinero, C.* (2005): Financial ratios, size, industry and interest rate issues in company failure: an extended multidimensional scaling analysis. *Kent Business School Working Papers 10*. University of Kent, Kent
- Virág M.* (1996): Pénzügyi elemzés, csődelőrejelzés. *Kosuth Kiadó*, Budapest
- Virág M. – Kristóf T.* (2006): Iparági rátákon alapuló csődelőrejelzés sokváltozós statisztikai módszerekkel, *Vezetéstudomány*, XXXVII. évf. 1. sz. 25–35. old.

Cikk beérkezett: 2008. 6. hó

Lektorai vélemény alapján véglegesítve: 2008. 7. hó

CIKKEK ANGOL NYELVŰ ÖSSZEFOGLALÓI

BARAKONYI, Károly

Role of the changes of the business schools – Challenges and changes in the management education and the MBAs

The American business schools at the early times functioned as trade schools, define business management as a profession. During the Cold War increasing the standards of education became a critical question of national defense: in the next decades the commercial schools transformed themselves into citadel of management science. But at the same time undesired effects sprang to existence: competing for scientific excellence, concentrating on research, fetishism of theory and methodology swap business schools too far from the business practice. Publication of article in theoretical periodicals, impact factors and references became more important than cooperation with practice. By now management education arrived to a new turning point: changes are necessary to create a new equilibrium between management as a science and management as a profession.

GELEI, Andrea

The business network – the quasi organization of the global economy

The aim of the paper is to present and interpret the basic building element of global business: the business network, its structure and operation. First basic terms – network, supply chain, supply network – are defined and described, than those changes are introduced that played significant role in increasing their importance. Characteristics of the new network economy are presented; especially changes in the coordination mechanism between cooperating parties in the network are demonstrated. Finally the two building blocks of global business networks: (i) nodes (business units) and (ii) threads (partnerships) are described in details.

FÜZY, Annamária

The new generation of knowledge management and the keys of its success

What is the secret ingredient of an information technology project supporting the learning of an organization? What are the current tools to enhance knowledge and

make it available for everyone? The 2.0 revolution has arrived already. Nowadays, the capability of learning has become the main source of the sustainable competitive advantages. Paralelly, the effective IT solutions have come forward as well. However, information technology can only influence the organization's competitiveness indirectly, by supporting the organizational learning. The aim of this article is therefore to introduce the new instruments of knowledge management, and show the key factors of their success.

VIRÁG, Miklós – KRISTÓF, Tamás

Multidimensional scaling in bankruptcy modeling

This article presents the applicability of multidimensional scaling techniques in the field of bankruptcy prediction based on empirical research. The relative neglect of multidimensional scaling techniques in bankruptcy literature does not coincide with the high level of their reliability. On the basis of empirical research it can be concluded that multidimensional scaling – in combination with logistic regression analysis – is competitive with other known bankruptcy modeling methods, as far as classification power and interpretability are concerned.

VILMÁNYI, Márton – Hetesi, Erzsébet

Could the performance of business relationships be modeled? – Literature summary and model development

This study analyses the measuring problems of performance of the inter-organisational relationships. In the first parts of this study the authors review the results of the relevant literature and present their suggested model which would be able to measure the relationship performance. In their research they proceed from a valuebased approach of buyer-supplier cooperation in which the relationship performance category is divided into economic directly evaluable and non-evaluable elements. The relationship performance is defined a multi dimensional category and the authors attempt to identify its most important factors. The relationship performance is described with the dimensions of results, the processes and the abilities. The model has been in a trial version, its empiric test was done, but the frames of this study does not allow of deep-analysis of its results.