

Metamódszerek alkalmazása a csődelőrejelzésben*

Virág Miklós – Nyitrai Tamás

A klasszifikációs feladatok megoldására jellemzően egy-egy kiválasztott módszert alkalmaznak a gyakorlatban. Ugyan a legkorszerűbb eljárások kimagasló találati arány elérésére képesek, a nemzetközi kutatási eredmények azt mutatják, hogy a gyengébb klasszifikációs teljesítményt mutató módszerek együttes alkalmazásával (ensemble) hasonlóan magas találati arány érhető el. A cikk fő célja a csődelőrejelzésben leggyakrabban alkalmazott két metamódszer (AdaBoost, bagging) előrejelző képességének összehasonlítása egy 976 hazai vállalkozás adataiból álló mintán. A cikk másik célkitűzése annak a vizsgálata, hogy az egyes pénzügyi mutatóknak a szakágazati átlagtól vett eltéréseire épített csődmodellek találati aránya hogyan viszonyul a nyers pénzügyi mutatókra, illetve az azok dinamikáját is figyelembe vevő változókra épített modellek találati arányához.

Journal of Economic Literature (JEL) kódok: C38, C49, G33

Kulcsszavak: csődelőrejelzés, metamódszerek, iparági átlag, döntési fák

1. Bevezetés

A gazdasági válság, valamint a fokozódó piaci verseny hatására az egész világgazdaságot egyre nagyobb bizonytalanság jellemzi, amely megnehezíti az üzleti döntések meghozatalát, emiatt a vállalatok jövőbeli fennmaradásával kapcsolatos előrejelzések fontossága minden eddiginél jelentősebb (Cao, 2012). A csődesemény jelentős veszteségeket okozhat a vállalati tevékenység valamennyi érintettjének, kimondottan a tulajdonosoknak, a vezetőségnek, de végső soron a teljes nemzetgazdaságnak is. Ebből adódóan a csődelőrejelzés az utóbbi időben a pénzügyi-számviteli szakirodalom egyik legfontosabb területe (Kim–Kang, 2012).

Módszertani szempontból a csődelőrejelzés egy egyszerű klasszifikációs feladat, amelynek célja a fizetőképés és a fizetéképtelen vállalkozások minél pontosabb megkülönböztetése. Az erre a célra alkalmas módszertani eljárások száma *Du Jardin* (2010) szerint meghaladja az 500-at, ennek ellenére nincs konszenzus arra vonatkozóan, hogy ezek közül melyik módszert célszerű alkalmazni a csődelőrejelző modellek építése során (Oreski et al., 2012).

* Jelen cikk a szerző nézeteit tartalmazza, és nem feltétlenül tükrözi a Magyar Nemzeti Bank hivatalos álláspontját. Virág Miklós, egyetemi tanár (Budapesti Corvinus Egyetem Vállalkozások Pénzügyei Tanszék); Nyitrai Tamás, doktorjelölt, tudományos segédmunkatárs (Budapesti Corvinus Egyetem Vállalkozások Pénzügyei Tanszék).

A csődelőrejelzés fejlődésére jelentős hatást gyakorol a klasszifikációs célokra alkalmas statisztikai és adatbányászati módszerek, valamint az azt támogató informatika fejlődése. Az utóbbi időszakban mindkét területet rohamos fejlődés jellemzi, ami nem hagyta érintetlenül a csődelőrejelzést sem, amelynek főáramát napjainkban is a módszertani összehasonlító elemzések bemutatása jellemzi (*Sánchez-Lasheras et al., 2012*). E területen napjainkban egy új kutatási irány a módszerkombinációk alkalmazási lehetőségeinek vizsgálata (Cao, 2012).

A módszerkombinációk egy speciális típusát metamódszereknek (ensemble) nevezik. Ezeknek az a lényege, hogy a modellépítés céljából rendelkezésre álló mintából többször véletlenszerűen visszatevéses mintát vesznek, és ezeken alkalmaznak egy előzetesen kiválasztott klasszifikációs eljárást. Ezt követően az egyes modellek által adott előrejelzéseket átlagolva alakul ki a modell végső találati aránya, amely jellemzően meghaladja az adott módszer önálló alkalmazásával elérhető találati arányt. A cikk egyik fő célja, hogy bemutassa és összehasonlítsa a nemzetközi szakirodalomban leginkább elterjedt két metamódszert: az AdaBoost és a bagging¹ eljárást; valamint az, hogy az ezek alkalmazásával elérhető találati arány meghaladja-e a tanulmányban választott klasszifikációs módszer önálló alkalmazásával elérhető pontosságot.

A csődelőrejelzésben hagyományosan a hányados típusú pénzügyi-számviteli mutatókat használják a modellek magyarázó változóiként. E mutatókat azonban a legtöbbször viszonyítási alap nélkül alkalmazzák a modellezés során, annak ellenére, hogy már az egyetemi tankönyvek is alapelveként tanítják: a pénzügyi mutatószámok nem jelentenek abszolút kritériumot, azokat csak valamilyen viszonyítási alap tükrében lehet objektíven megítélni (Virág et al., 2013). E célra *Platt–Platt* (1990) az iparági relatív mutatók alkalmazását javasolta, ami a vállalati mutató értékét az iparági átlag 100-szorosához viszonyítja. E megközelítés előnye, hogy összehasonlíthatóvá teszi a különböző iparágakban tevékenykedő vállalatok pénzügyi mutatóit, illetve megoldást nyújt a csődmodellek időbeli instabilitásának² problémájára is.

Az előnyös tulajdonságok mellett érdemes szót ejteni az iparági relatív mutatók alkalmazásakor felmerülő nehézségekről is. Fontos figyelembe venni, hogy több pénzügyi mutatószámot intervallumskálán mérnek, ami az iparági relatív mutatók esetén problémát okozhat. Félrevezető lehet ugyanis, ha egy negatív vállalati jövedelmezőséget egy szintén negatív iparági átlaghoz viszonyítunk.³ Ekkor ugyanis az iparági relatív mutató pozitív értéke nem szükségképpen mutat reális képet az iparági átlaghoz képest. E problémát kiküszöbölendő,

1 Tudomásunk szerint az AdaBoost (adaptive boosting) és a bagging (bootstrap aggregating) eljárásokra vonatkozóan nem áll rendelkezésre pontos magyar megfelelő, így a továbbiakban is az angol megnevezéseket használjuk a módszerekre történő utalások során.

2 A csődmodellek időbeli instabilitása azt jelenti, hogy egy adott időpont, vagy időszak adatain felállított modellek találati aránya a későbbi adatokon jelentősen csökken.

3 Negatív iparági átlag a rendelkezésre álló adatbázisunk szerint a recessziós gazdasági helyzetben több iparágban is előfordult.

a tanulmány a vállalati mutatóknak a saját szakágazati átlaguktól vett különbségét használja fel magyarázó változóként azok hányadosa helyett. A módszertani összehasonlítás mellett azt is vizsgáljuk, hogy a szakágazati átlagtól számított eltérések – mint magyarázó változók – milyen találati arány elérését teszik lehetővé egy 976 hazai vállalkozás adatait tartalmazó minta esetén.

A nemzetközi szakirodalom napjainkban szintén egyre intenzívebben foglalkozik azzal a problémával, amely abból ered, hogy modellezők döntő többsége csak a csőd előtti legutolsó év pénzügyi mutatóit használja fel a modellek magyarázó változói között. E megközelítés problémája, hogy figyelmen kívül hagyja a vállalati gazdálkodás folyamatjellegét, s kizárólag a vállalatok statikus pénzügyi helyzetét veszi alapul a modellezés során (*Chen et al., 2013*). A cikk feltételezése szerint releváns információt hordozhat az, hogyan viszonyul egy vállalat pénzügyi mutatóinak egy-egy konkrét értéke az azt megelőző időszak megfelelő értékeihez. A cikk a nyers pénzügyi mutatók, valamint azoknak a szakágazati átlagtól vett eltérései mellett vizsgálja azt is, hogy milyen találati arány érhető el abban az esetben, ha külön változóként szerepeltetjük a csődmodellekben azt is, hogyan viszonyul az egyes pénzügyi mutatók értéke a megelőző időszak azonos változóinak értékeihez.

A tanulmány kísérletet tesz e három változókör (nyers, szakágazati átlagtól vett eltérés, dinamikus) önálló, illetve kombinált alkalmazására a csődelőrejelző modellekben. A vizsgálat célja annak a megállapítása, hogy ezek a változócsoportok együtt vagy külön alkalmasak-e inkább a csődmodellek találati arányának maximalizálására, ami a csődmodellek fejlesztésének elsődleges célja (*Du Jardin, 2010*).

Fontosnak tartjuk hangsúlyozni, hogy a cikkben bemutatott elemzés kísérleti jellegű. Nem reprezentatív mintára épül, és nem célja egy konkrét előrejelzésre optimalizált csődmodell prezentálása. Véleményünk szerint azonban a rendelkezésre álló minta és az azon elért eredmények alkalmasak lehetnek a feltett kutatási kérdések megválaszolására, valamint az eredményekből levonható következtetések további kutatások tárgyát képezhetik a csődelőrejelzésben és az attól független kutatási területeken is.

A tanulmány felépítése a következő: a második szakaszban olyan nemzetközi publikációkat mutatunk be, amelyek összehasonlító elemzéssel vizsgálták a metamódszerek teljesítményét a csődelőrejelzésben. A harmadik szakaszban röviden bemutatásra kerül az AdaBoost és a bagging eljárás. A negyedik szakasz az empirikus vizsgálathoz felhasznált mintát és az alkalmazott kutatási módszereket ismerteti. Az ötödik szakaszban olvashatók a vizsgálatok eredményei. A hatodik, záró szakaszban összefoglaljuk az elemzésből levonható következtetéseket, illetve az azokból adódó, lehetséges további kutatási irányokat.

2. A metamódszerek elméleti alapjai

A metamódszerek működését *Marqués et al. (2012a)* alapján ismertetjük. Ezeket az eljárásokat szó szerint lehet módszeregyütteseknek (ensemble) is fordítani, mivel egy adott klasszifikációs módszert többször is alkalmazunk egy adathalmazon. A metamódszerek alkalmazása során vagy a megfigyelések, vagy a rendelkezésre álló input változók köréből veszünk egy visszatevéses mintát, és a klasszifikációs módszert csak a kiválasztott részmintán alkalmazzuk.

A metamódszerek alkalmazásakor előre rögzíteni kell, hogy hányszor veszünk mintát a rendelkezésre álló adatkörből. P darab részminta esetén az eljárást p -szer futtatjuk. Az eredmények egyenlő súllyal jelennek meg a metamódszerek végső előrejelzésének kialakításában; más szóval, az egyes megfigyeléseket abba a csoportba soroljuk, amelybe a p futtatás során a legtöbbször kerültek.

Ahhoz, hogy a metamódszerekkel elérhető találati arány meghaladja a módszerek egyedi alkalmazásával elérhető eredményeket, szükséges, hogy az egyes modellfuttatások eredményei a lehető legnagyobb mértékben különbözzenek. Ez a diverzitás elérhető úgy, hogy a p futtatás mindegyikében más részmintát veszünk alapul a rendelkezésre álló megfigyelések közül. Ezt az elvet követi az AdaBoost és a bagging eljárás is. Utóbbi p darab visszatevéses mintát vesz a rendelkezésre álló adathalmazból, és azok mindegyikén futtatja a felhasználó által választott klasszifikációs módszert, majd azok eredményeinek egyszerű számtani átlagaként alakítja ki a végső klasszifikációt. Az AdaBoost eljárás ettől csak abban különbözik, hogy a mintavétel a modellfuttatások eredményeinek függvényében változik oly módon, hogy azok a megfigyelések, amelyek az i -edik futtatáskor tévesen kerültek besorolásra, az $i+1$ -edik mintába nagyobb eséllyel kerülnek be, míg azok, amelyek az i -edik részminta alapján a megfelelő osztályba soroltattak, kisebb eséllyel jelennek meg az $i+1$ -edik részmintában. A végső klasszifikáció ebben az esetben is abba az osztályba történik, ahová az egyes megfigyelések a p futtatás többségében kerültek.

A klasszifikációs módszerek teljesítménye érzékeny a modellépítéshez felhasznált független változók relevanciájára is. Különböző input változók felhasználásával általában különböző találati arány érhető el azonos klasszifikációs módszerek alkalmazásakor. A metamódszerek hatékony működéséhez szükséges diverzitás akkor is biztosítható, ha nem a rendelkezésre álló megfigyelésekből, hanem a rendelkezésre álló változók köréből veszünk visszatevéses mintát. Ez a módszer véletlen alterek (*random subspace*) néven ismert a szakirodalomban.

Jelen tanulmányban az első megközelítést alkalmazzuk: azt az esetet, amikor az együttesen belüli diverzitást a megfigyelésekből vett visszatevéses minták alkalmazása biztosítja. A cikk a korábban bemutatott AdaBoost és bagging módszert hasonlítja össze hazai vállalkozások adatain, amelyről a szerzők tudomása szerint a gazdasági vonatkozású, hazai szakirodalomban első alkalommal készült tudományos publikáció.

3. Szakirodalmi áttekintés

A metamódszerek alkalmazása napjainkban is aktuális kutatási kérdés a csődelőrejelzés nemzetközi szakirodalmában. A következőkben az utóbbi években megjelent összehasonlító elemzések eredményei közül ismertetünk néhányat a teljesség igénye nélkül.

Alfaro et al. (2008) spanyol vállalatok adatain vetette össze a neurális hálók és a döntési fákon alkalmazott AdaBoost módszer előrejelző képességét. Az idézett szerzők eredményei szerint a meta eljárás alkalmazásával felállított modell eredménye szignifikánsan meghaladta a neurális hálók és a lineáris diszkriminancia-analízis teljesítményét is.

Ezzel szemben orosz termelő vállalatok adatain *Fedorova et al. (2013)* nem talált érdemi különbséget a neurális hálók önálló, illetve az AdaBoost eljárásban történő együttes alkalmazása között.

A metamódszerekben egy adott klasszifikációs módszert többször alkalmaznak. A klasszifikációs módszerek teljesítményét azonban jelentősen befolyásolja a felhasznált magyarázó változók minősége, ami a változószelekció fontosságára hívja fel a figyelmet a csődelőrejelzésben. *Wang et al. (2014)* szerint a csődmodellek pontosságát számottevően ronthatja a felesleges magyarázó változók jelenléte az input változók között. Ebből adódóan az idézett szerzők kísérletet tettek egy változószelekciós eljárás integrálására az AdaBoost eljárásba. A komplex módszerrel felállított csődmodellek találati aránya szignifikánsan meghaladta a hagyományos metamódszerekkel, illetve a leggyakrabban használt egyedi technikákkal elért eredményeket.

A csődelőrejelzés témakörében napjainkban a legtöbb tanulmány a különböző módszerek összehasonlító elemzését tárgyalja (*Sánchez-Lasheras et al., 2012*). Ennek alapját a klasszifikációs célra alkalmas adatbányászati módszerek és az azt támogató informatika fejlődése teremti meg; motivációja pedig elsősorban abból ered, hogy máig nincs konszenzus abban, melyik módszert érdemes a leginkább alkalmazni a csődelőrejelzésben (*Oreski et al., 2012*). Az egyedi módszerek összehasonlító elemzéseinek eredményei arra utalnak, hogy egyes esetekben egyik, míg más esetekben egy másik módszer alkalmazásával érhető el a legmagasabb találati arány. Így joggal merül fel a kérdés, hogy a metamódszerek esetén is hasonló tendencia érvényesül-e. A választ *Marqués et al. (2012a)* próbálta megadni; tanulmányukban hét klasszifikációs módszer teljesítményét hasonlították össze a leggyakrabban alkalmazott metamódszerek esetén, hat csődmodell adatbázisa alapján. Eredményeik arra engedtek következtetni, hogy a legmegbízhatóbb eredményt a C4.5 klasszifikációs eljárás mutatja a metamódszerekben. Ebből adódóan cikkünk is ezt a technikát használja.

Marqués et al. (2012b) arra a kérdésre kereste a választ, hogy javít-e a modellek előrejelző képességén a metamódszerek két megközelítésének (a megfigyelésekből, illetve a változókból történő mintavételnek) az együttes alkalmazása. Elemzéseiket hat különböző csődmodell adatbázisán végezték. Az eredmények arra utalnak, hogy a metamódszerek

együttes alkalmazása több esetben javított az egyedi eljárásokkal, illetve a metamódszerek önálló alkalmazásával elért eredményeken.

A metamódszerek egy speciális területének tekinthető az, amikor nem egy adott módszert alkalmaznak többször a tanuló adathalmazban lévő megfigyelések és/vagy változók egy részhalmazán, hanem a teljes adatkörön több különböző klasszifikációs eljárást futtatnak, és ezek eredményeit kombinálva adódik a megfigyelések végső besorolása az egyes csoportokba. E területen Cao (2012) munkája említendő, aki a kínai tőzsdén jegyzett vállalatok adatain tett kísérletet a különböző módszerek eredményeinek ilyen kombinációjára. Az idézett szerző eredményei ugyan szignifikánsan meghaladták az egyedi módszerekkel elérhető találati arányokat, hozzá kell azonban tenni, hogy a hagyományos metamódszerekkel szemben nem végzett összehasonlításokat, emiatt pedig nem támasztja alá empirikus kutatási eredmény, hogy az egy adott klasszifikációs módszert többször alkalmazó metamódszerek vagy a több klasszifikációs módszer eredményeit integráló eljárások tekinthetők-e jobbnak a csődelőrejelzés területén.

4. A vizsgált adathalmaz és az alkalmazott kutatási módszerek

4.1. Minta

A feltett kutatási kérdések megválaszolása céljából saját adatgyűjtésből összeállítottunk egy 976 elemű mintát, amely 51-49 százalékos arányban tartalmaz fizetőképessé és fizetéstelen vállalatokat. Ebből adódóan a minta nyilvánvalóan nem tekinthető reprezentatívnak, ami általánosnak mondható a csődelőrejelzésben. A csődös cégek felülreprezentálása azzal magyarázható, hogy a gépi tanulásra épülő adatbányászati eljárások egyenlőtlen megoszlás esetén hajlamosak a domináns csoport sajátosságaira specializálódni (Horta–Camanho, 2013), ami a csődelőrejelzésben a fizetéstelen vállalkozások túlságosan alacsony találati arányát eredményezheti. A mintavétel során azok a vállalkozások minősültek fizetéstelennek, amelyek az adatgyűjtés időpontjában a Cégjegyzék adatai szerint csőd- vagy felszámolási eljárás alatt álltak.

A mintavétel során érvényesített szempontok:

1. Csak olyan megfigyelések kerültek be a mintába, amelyek adatai legalább három évre visszamenőleg hozzáférhetőek voltak a Közigazgatási és Igazságügyi Minisztérium Céginformációs és az Elektronikus Cégeljárásban Közreműködő Szolgálatának internetes

oldalán.⁴ E mintavételi szempont alkalmazását az indokolja, hogy lehetőség szerint kiszűrjük a nagyon fiatal vállalkozásokat, amelyek a kezdeti nehézségek miatt gyakran jobban hasonlítanak az idősebb, de fizetésképtelen vállalkozásokra, mint a működőkre (Du Jardin, 2010). Másrészt a tanulmányban olyan dinamikus pénzügyi mutatókat is alkalmazunk, amelyek a vállalatok pénzügyi mutatóinak legutoljára megfigyelt értékét az azt megelőző időszak megfelelő adataihoz viszonyítják, ehhez pedig szükséges, hogy legalább három évre visszamenőleg rendelkezésre álljanak a pénzügyi mutatók értékei.

2. Kimaradtak a mintából az olyan megfigyelések, amelyek legalább két egymást követő évben nem realizáltak árbevételt. Ennek oka, hogy az ilyen vállalkozások vélhetően nem folytatnak érdemi gazdálkodást, így mintába kerülésüknek torzító hatása lenne a modellek eredményeire.
3. Nem kerülhettek be a mintába olyan megfigyelések, amelyek pénzügyi mutatói a vizsgált, legalább hároméves időtávon nem mutattak szóródást. Ebben az esetben ugyanis a tanulmányban alkalmazott, dinamikus pénzügyi mutatók számításához felhasznált, standardizált értékeket nem lehet kiszámítani, ugyanis azokhoz a megfigyelt adatok szóródását használtuk viszonyítási alapként.

Du Jardin (2010) szerint a csődelőrejelzésben gyakran alkalmazott megközelítés olyan pénzügyi mutatók használata magyarázó változóként, amelyek más tanulmányokban eredményesnek bizonyultak. E tanulmány is ezt a megközelítést alkalmazza. Az input változók kiválasztása során az első hazai csődmodell változóit (részletesen I. Virág–Hajdu, 1996) és saját megfontolásainkat vettük figyelembe. A kiszámított 17 mutató nevét és számítás módját mutatja az 1. táblázat. A mutatószámok tartalmát részletesen Virág et al. (2013) tárgyalja. Az egyes mutatók számítása során az egyes mérlegtételeket, illetve az eredménykimutatás érintett sorait azok fordulónapi záró értékén vettük figyelembe.

A csődmodellekben gyakran használt mutatószámok közé tartozik a sajáttőke-arányos nyereség, amely gyakran veti fel a kettős negatív osztás problémáját (Kristóf 2008). Ennek a kezelésére nincs egyértelműen preferált megoldás a szakirodalomban, ezért ezt a mutatót nem vettük figyelembe a számítások során.

A hányados típusú mutatószámok másik jellemző problémája akkor merül fel, amikor a hányados nevezőjében nulla érték adódik. E problémát a gyakorlatban gyakran kezelik úgy, hogy az ilyen adatokat hiányzó értéknek tekintik, majd a többi megfigyelés valamilyen középértékével, vagy azok valamelyik szélső percentilisével helyettesítik. Véleményünk szerint azonban az előbbi megközelítés nem feltétlenül visz konzisztens értéket a csődelőrejelző modellekbe, míg utóbbi esetben mintaszpecifikus lehet, hogy az adott értéket mivel helyettesítjük. Munkánkban azt a megközelítést alkalmaztuk, hogy azokban az esetekben, ahol a mutató nevezője nulla értéket venne fel, ott a nevező értékét 1-nek választottuk.

4 <http://e-beszamolo.kim.gov.hu/kereses-Default.aspx>

1. táblázat**Az empirikus vizsgálatban felhasznált mutatószámok neve és számítás módja**

Mutatószám neve	Számításmódja
Likviditási ráta	Forgóeszközök/rövid lejáratú kötelezettségek
Likviditási gyorsráta	(Forgóeszközök-készletek)/rövid lejáratú kötelezettségek
Pénzeszközök aránya	Pénzeszközök/forgóeszközök
Cash flow/kötelezettségek	(Adózás utáni eredmény + értékcsökkenési leírás)/kötelezettségek
Cash flow/rövid lejáratú kötelezettségek	(Adózás utáni eredmény + értékcsökkenési leírás)/rövid lejáratú kötelezettségek
Tőkeellátottság	(Befektetett eszközök + készletek/saját tőke
Eszközök forgási sebessége	Értékesítés nettó árbevétele/mérlegfőösszeg
Készletek forgási sebessége	Értékesítés nettó árbevétele/készletek
Követelések forgási ideje	Követelések/értékesítés nettó árbevétele
Eladósodottság	Kötelezettségek/mérlegfőösszeg
Saját tőke aránya	Saját tőke/mérlegfőösszeg
Bonitás	Kötelezettségek/saját tőke
Árbevétel-arányos nyereség	Adózás utáni eredmény/Értékesítés nettó árbevétele
Eszközarányos nyereség	Adózás utáni eredmény/mérlegfőösszeg
Követelések/rövid lejáratú kötelezettségek	Követelések/Rövid lejáratú kötelezettségek
Nettó forgótőke aránya	(Forgóeszközök-rövid lejáratú kötelezettségek)/Mérlegfőösszeg
Vállalat mérete	Az eszközállomány természetes alapú logaritmus

A fenti szempontok alapján végrehajtott mintavétel eredményeképpen rendelkezésre áll egy 976 hazai vállalkozás pénzügyi mutatóiból, valamint azok szakágazati átlagtól⁵ vett eltéréseit tartalmazó adatbázis a 2001–2012-es időszakra. A mintába került megfigyelések legutolsó megfigyelt üzleti éve a 2009–2012-es időintervallumba esik. A minta véletlen jellegéből adódóan rendkívül heterogén: egyaránt tartalmaz meglehetősen alacsony eszközállománnyal jellemezhető mikrovállalkozásokat, valamint közepes- és nagyvállalatokat. A tevékenységi kör tekintetében hasonlóan heterogén a minta, amelyben a gazdaság valamennyi fontosabb részterülete képviselteti magát (mezőgazdaság, ipar, kereskedelem, informatika stb.). A nagyfokú heterogenitás ellenére a cikkben bemutatott modellek 80% körüli találati arányt mutattak, ami véleményünk szerint alapul szolgálhat a tanulmány eredményeiből levonható következtetések gyakorlati alkalmazására és reményeink szerint további kutatásokat ösztönöz majd.

5 A mintába került vállalkozások szakágazati besorolását a Cégjegyzék elektronikusan szabadon hozzáférhető, a mintavétel időpontjában hatályos adatai alapján határoztuk meg.

4.2. Kutatási módszerek

A tanulmány elsődleges célja az AdaBoost és a bagging eljárás előrejelző teljesítményének összevetése a C4.5 klasszifikációs módszer alkalmazásával, amely egy döntési fát generáló adatbányászati eljárás. A metamódszerek tagszámát (p) 100-nak választottuk, a C4.5 eljárás alkalmazása során legalább 5 megfigyelés volt szükséges egy újabb ág képzéséhez. A felállított döntési fák végső ágain a besorolás abba a csoportba történt, ahol az adott csoport (fizetőképes/fizetéképtelen) aránya magasabb volt.⁶ A klasszifikációs technika módszertani háttérét röviden Nyitrai (2014) ismerteti, részletesen pedig *Quinlan* (1993) munkáját ajánljuk az olvasó figyelmébe.

A vizsgált módszerek előrejelző képességét a minta 100-szoros véletlenszerű felosztásával becsültük. Az eljárás lényege, hogy a rendelkezésre álló adathalmaz 75-25%-os arányban felosztásra került tanuló és tesztelő mintára 100 darab véletlenszerűen kijelölt osztópont segítségével. Az előrejelző teljesítményt a száz darab tesztelő minta találati arányának átlagával mérjük. A modellek találati aránya alatt a helyesen besorolt megfigyelések arányát értjük az összes megfigyeléshez viszonyítva.

A cikk másik célja annak a vizsgálata, hogy a nyers pénzügyi mutatók mellett, illetve helyett a modellek független változói között szerepeltetjük a pénzügyi mutatók iparági átlagtól vett különbségét, illetve olyan dinamikus pénzügyi mutatókat is, amelyek azt számszerűsítik, hogy egy legutoljára megfigyelt év pénzügyi mutatója hogyan viszonyul az adott vállalkozás korábbi megfelelő pénzügyi mutatójának értékeihez. Ezt az alábbi formula alkalmazásával számszerűsítettük:

$$\frac{X_{i,t-1} - X_{i,\min_{(t-2:t-n)}}}{X_{i,\max_{(t-2:t-n)}} - X_{i,\min_{(t-2:t-n)}}$$

A képletben i egy konkrét vállalkozást, t azt az évet, amelyre vonatkozóan előrejelzést kívánunk adni, n pedig az adott megfigyeléshez rendelkezésre álló idősor hosszát (a megfigyelt évek számát) jelöli.

Mivel a megfigyelt vállalatok mutatószám-idősorai számos esetben tartalmaztak kiugró értékeket, minden megfigyelés mindegyik pénzügyi mutatószám-idősorát standardizáltuk az idősor átlagával és szórásával⁷, majd azokat az értékeket, amelyek a két szórássterjedelmen kívülre estek, a hozzájuk legközelebb eső, de még a terjedelmen belül lévő értékekkel helyettesítettük⁸ annak érdekében, hogy a fenti formula minél pontosabban kifejezze, ho-

6 Ezt a megközelítést csak abban az esetben célszerű alkalmazni, amikor a mintában a két csoport 50-50 százalékos arányban van jelen. Eltérő esetben a modellek teljesítményét célszerű a Gini-mutató vagy a ROC-görbe segítségével megítélni, elkerülve a cut-off érték szubjektív meghatározását.

7 A standardizáláshoz felhasznált átlag és szórás számításánál az egyes megfigyelések mutatószám-idősorainak a $t-2$ és $t-n$ időpontok közötti adatait használtuk fel.

8 A legutoljára megfigyelt ($t-1$) év adatát nem helyettesítettük és nem használtuk fel helyettesítésre annak érdekében, hogy a vizsgált vállalkozás pénzügyi helyzetét a legutoljára megfigyelt évben a lehető legpontosabban megítélhessük az azt megelőző időszak tükrében.

gyan viszonyul a vizsgált vállalkozás legutóbb megfigyelt mutatója az azt megelőző időszak megfelelő értékeihez.

Felmerülhet a kérdés, hogy a mutatószám-idősorok esetén miért a 2 szórásterjedelem alkalmazásával definiáltuk az outliereket. Statisztikai hüvelykujjszabály az outlierok szűrése során az 5, 3, illetve 2 szórásterjedelem alkalmazása. Ezek közül azért esett a választásunk a „legszigorúbb” értékre (2 szórás), mert az egyes megfigyelések esetén az átlag és szórás számítása céljából rendelkezésre álló idősor 2–11 elemű. Kísérleti számításaink azt mutatták, hogy ilyen rövid idősorok esetén a dinamikus mutatókra épített csődmodellek jobb előrejelző teljesítményt mutatnak a „szigorúbb” szabályok alkalmazásával.

5. Az empirikus vizsgálatok eredményei

A tanulmányban bemutatott kutatás két fő célt tűzött maga elé:

1. Megvizsgálni, hogy mutatkozik-e szignifikáns különbség az AdaBoost és a bagging eljárás találati aránya között a C4.5 klasszifikációs módszer alkalmazása esetén. Itt felmerül az a kérdés is, hogy érdemes-e egyáltalán a metamódszereket alkalmazni a C4.5 eljárás önálló alkalmazása helyett. Ennek vizsgálata céljából a számításokat elvégeztük a választott klasszifikációs módszer önálló alkalmazásával is.
2. Összevetni a nyers pénzügyi mutatók, az azokból képzett dinamikus ráták, illetve a nyers mutatók szakágazati átlagtól vett eltéréseinek alkalmazhatóságát a fenti modellek input változóinak szerepében.

A felsorolt kutatási kérdésekre az előző szakaszban bemutatott kutatási módszerek alkalmazásával kerestünk választ. Számításaink eredményeit a 2. táblázat foglalja össze.

2. táblázat								
A kísérleti modellfuttatások találati arányai a tesztelő minták átlagában								
Módszer	Input változókör							Átlag
	nyers	iparági	dinamikus	nyers dinamikus	nyers iparági	nyers dinamikus iparági	dinamikus iparági	
AdaBoost	78,80%	76,06%	78,67%	81,33%	79,33%	81,51%	80,34%	79,43%
Bagging	79,88%	77,97%	79,91%	83,01%	79,76%	82,17%	81,29%	80,57%
Önálló	74,34%	70,64%	72,37%	75,12%	73,29%	74,91%	73,31%	73,43%
Átlag	77,67%	74,89%	76,98%	79,82%	77,46%	79,53%	78,31%	

Az eredmények alapján az alábbi következtetések vonhatók le:

- A bagging eljárás találati aránya valamennyi vizsgált esetben meghaladta az AdaBoost eljárással kapott eredményt, meg kell azonban jegyezni, hogy a különbség jellemzően 1 százalékpont körüli.
- A C4.5 módszer önálló alkalmazása, valamint a bagging eljárás esetén a legjobb teljesítmény abban az esetben adódott, amikor a nyers mutatók mellett szerepeltek az azokból képzett dinamikus mutatók is; viszont az AdaBoost módszer esetén a legjobb előrejelző teljesítményt az a modell mutatta, amelyben mindhárom változócsoporthoz szerepelt a független változók körében. A felállított modellek közül a legjobb előrejelző teljesítményt akkor kaptuk, amikor a bagging módszer alkalmazásakor a nyers mutatók mellett a magyarázó változók körében szerepeltek a dinamikus mutatók is.

Kérdés azonban, hogy a fenti eredmények mennyiben tudhatók be a mintavételi sajátosságoknak, azaz mennyire tekinthető szignifikánsnak a találati arányok különbözősége az egyes módszerek esetén. Tekintettel arra, hogy számításaink alapján a három vizsgált eljárás egyike esetén sem feltételezhető a találati arányok normális eloszlása, a besorolási pontosságok szignifikáns különbözőségét *Mann* és *Whitney* próbájával vizsgáltuk, ami a független mintás *t*-próba nemparaméteres megfelelője (Du Jardin, 2010). A teszt eredménye mindhárom összehasonlításban szignifikáns különbséget jelzett az összes releváns szignifikanciaszinten, ami azt jelenti, hogy a 2. táblázatban közölt eltérések szignifikánsak. Tehát a bagging eljárás találati aránya a C4.5 eljárás esetén, ha csak kis mértékben is, de szignifikánsan meghaladja az AdaBoost eljárás pontosságát; valamint mindkét metamódszer teljesítménye szignifikánsan felülmúlja a C4.5 klasszifikációs eljárás önálló alkalmazásával elérhető eredményt. Utóbbi esetben a különbség jóval markánsabb – 6-7 százalékpont körüli a differencia.

A tanulmány másik célkitűzése a különböző változókörok felhasználásával felállított modellek találati arányának összevetése volt. A 2. táblázat eredményei arra utalnak, hogy a bagging eljárás és a C4.5 önálló alkalmazása esetén a legjobb előrejelző teljesítmény akkor adódik, amikor a nyers pénzügyi mutatók mellett szerepeltetjük az azokból képzett dinamikus mutatókat is. Kivételt képez azonban az AdaBoost eljárás, ahol akkor mutatkozott a legjobb előrejelző teljesítmény, amikor mindhárom változócsoporthoz szerepelt a modellek független változóinak körében. Az egyes változóhalmazok alkalmazásával kapott találati arányok közti különbség szignifikanciájának tesztelése ebben az esetben is igényli azok normalitásának ellenőrzését. Az erre irányuló próba minden releváns szignifikanciaszinten elveti a találati arányok normális eloszlásának feltevését, így ebben az esetben is nemparaméteres teszthez kell folyamodnunk az egyes változókörok alkalmazásával kapott eredmények szignifikáns különbözőségének vizsgálata érdekében. A különböző független változóhalmazok közt páronként végrehajtott Mann–Whitney-próba *p*-értékeit mutatja a 3. táblázat.

3. táblázat

A vizsgált változócsoportok átlagos találati arányain végzett

Mann–Whitney-próba p -értékei

	nyers	iparági	dinamikus	nyers dinamikus	nyers iparági	nyers iparági dinamikus	iparági dinamikus
nyers	-	0,000	0,134	0,000	0,827	0,000	0,003
iparági		-	0,000	0,000	0,000	0,000	0,000
dinamikus			-	0,000	0,220	0,000	0,000
nyers dinamikus				-	0,000	0,389	0,000
nyers iparági					-	0,000	0,002
nyers iparági dinamikus						-	0,001
iparági dinamikus							-

Azokban az esetekben, ahol a p -érték közel áll a nullához, ott nagy valószínűséggel feltehető, hogy a 2. táblázatban közölt átlagos találati arányok nem csak a mintavételi sajátosságok miatt különböznek egymástól. Ez az esetek többségében teljesül. A tesztek eredményei a következőképp értelmezhetők:

Nem mutatkozott szignifikáns eltérés

- a nyers pénzügyi mutatók és a dinamikus ráták között;
- a nyers pénzügyi mutatók és azon eset között, amikor azok mellett szerepeltek az iparági átlagtól vett eltérések;
- a dinamikus mutatók és azon eset között, amikor a nyers pénzügyi ráták mellett szerepeltek az iparági átlagtól vett különbségek is;
- abban az esetben, amikor mindhárom változókört és azt az esetet vetettük egybe, amikor a nyers pénzügyi mutatók mellett szerepeltek az iparági átlagtól vett eltérések.

A 2. táblázatban közölt átlagok közti különbségek az előző esetek kivételével szignifikánsak. Ennek alapján statisztikailag is alátámasztott, hogy a legjobb előrejelző teljesítmény akkor adódott, amikor a nyers pénzügyi mutatók mellett az input változók körében szerepeltettük az azokból képzett dinamikus rátákat is. Ettől a teljesítménytől nem tér el szignifikánsan az az eset, amikor az iparági átlagtól való eltéréseket is szerepeltetjük a független változók között. Ez azonban arra utal, hogy utóbbiaknak nincs hozzáadott értéke a modellek előrejelző képessége szempontjából. Meglepő eredmény, hogy az iparági átlagtól vett eltérésekkel elérhető találati arány szignifikánsan a leggyengébb az összes lehetséges kombinációban.

6. Összegzés

Tanulmányunk elsődleges motivációját az a csődelőrejelzés nemzetközi szakirodalmában tapasztalható trend jelentette, amely a klasszifikációs módszerek önálló alkalmazása helyett a metamódszerek használatát javasolja. Kutatási eredmények azt mutatják, hogy ezzel a modellek találati aránya szignifikánsan javítható.

Tekintettel arra, hogy tudomásunk szerint a gazdasági vonatkozású, hazai szakirodalomban hasonló kutatás még nem készült, a két leggyakrabban használt metamódszer (AdaBoost, bagging) teljesítményét vetettük össze a C4.5 klasszifikációs módszer alkalmazása mellett. Választásunkat az indokolta, hogy Marqués et al. (2012a) összehasonlító elemzésének eredménye alapján ezen eljárás alkalmazása esetén érhető el a legnagyobb mértékű növekedés a klasszifikációs teljesítményben a módszer önálló alkalmazásával szemben.

Kutatásunk eredményei alapján is szignifikáns javulás mutatkozik a C4.5 eljárás esetén a metamódszerek alkalmazásával a módszer önálló használatával összevetve, függetlenül attól, hogy az AdaBoost vagy a bagging eljárást tekintjük.

Munkánkban összehasonlítottunk három változócsoporthat: a nyers pénzügyi mutatókat, az azokból képzett dinamikus rátákat, amelyek azt számszerűsítik, hogy hogyan viszonyul az egyes vállalkozások legutoljára megfigyelt mutatója az azt megelőző időszak megfelelő értékeihez, valamint azt, hogy a nyers pénzügyi mutatók értéke mekkora eltérést mutat az adott szakágazat átlagához képest. Az iparági mutatók esetében eltértünk a Platt–Platt (1990) által javasolt iparági relatív ráták formulájától, ugyanis annak értelmezése az intervallumskálán mért pénzügyi mutatók esetén problematikus. Kutatási eredményeink azt mutatták, hogy a vállalati mutatók szakágazati átlagtól vett eltéréseire épített modellek előrejelző képessége szignifikánsan elmarad a nyers és a dinamikus pénzügyi mutatókkal elérhető találati arányhoz képest a tanulmányban vizsgált összes lehetséges kombinációban.

A legjobb előrejelző teljesítményt akkor tapasztaltuk, amikor a nyers pénzügyi mutatók mellett az azokból képzett dinamikus rátákat is szerepeltettük a modellek input változói között. Ez a két változó kör közti szinergikus hatás jelenlétére utaló eredmény. A vizsgált változócsoporthat átlagos találati arányai között elvégzett nemparaméteres próba eredményei arra utalnak, hogy az e két változó csoport együttes alkalmazásával adódott, átlagos előrejelző képesség szignifikánsan felülmúlja a vizsgált három változó kör bármilyen kombinációjú alkalmazásával elérhető eredményt. Ez alól az egyedüli kivétel az volt, amikor mindhárom változó csoportot egyszerre alkalmaztuk, az így elérhető pontosság viszont nem különbözik szignifikánsan a nyers mutatókkal és az azokból képzett dinamikus rátákkal elérhető eredménytől – ez arra utal, hogy az iparági eltéréseknek nincs hozzáadott értéke a modell teljesítmény szempontjából.

Fontosnak tartjuk hangsúlyozni, hogy következtetéseinket kizárólag a C4.5 módszer alkalmazásával kapott eredmények alapján vontuk le. Az eredmények robusztusságát célszerű lehet a későbbiekben más klasszifikációs módszerek alkalmazásával is megvizsgálni.

Esetünkben ugyan a szakágazati átlagtól való eltérések nagyságára épített modellek nem hoztak érdemi javulást a nyers pénzügyi mutatókra épített modellekhez képest, az eredményből mégsem vonható le az a következtetés, hogy az iparági középértékek használata indokolatlan lenne a csődelőjelzésben. Hazai viszonyok között például Kristóf (2005) az előrejelző teljesítmény javulását tapasztalta az első hazai csődmodell adatbázisán végzett kutatásai során. Az idézett szerző a javulást részben az iparági ráták alkalmazásával indokolta. Ebből adódóan egy lehetséges jövőbeli kutatási irány lehet annak a vizsgálata, hogyan lehet kezelni az iparági relatív mutatók értelmezési problémáját az intervallumskálán mért pénzügyi mutatók esetén oly módon, hogy az iparági középértékhez viszonyított pénzügyi mutatók javítsák a modellek előrejelző képességét. A cikk eredményei alapján érdekes kutatási terület lehet annak a vizsgálata, hogy milyen klasszifikációs teljesítmény érhető el az iparági rátákból kalkulálható, dinamikus pénzügyi mutatók használatával.

További vizsgálatot igénylő kérdés, hogy az iparági középértékekhez viszonyított pénzügyi mutatók alkalmazásával miért nem javult a modellek előrejelző képessége. Véleményünk szerint erre egy lehetséges magyarázat az lehet, hogy a mintában szereplő vállalkozások tevékenységi körét a Cégjegyzékben hatályosan bejegyzett főtevékenységi TEÁOR-kód alapján határoztuk meg. Elképzelhető, hogy a megfigyelt vállalkozások gazdálkodásában jelentős szerepet játszanak egyéb melléktevékenységek, amelyekről a nyilvánosan hozzáférhető adatbázisok használatakor nem volt információnk. Feltételezhető, hogy e melléktevékenységek jelentősége oly mértékű, hogy a főtevékenységi TEÁOR-kód alapján azonosított szakágazati átlagok kevésbé alkalmasak arra, hogy a pénzügyi mutatók esetén viszonyítási alapként használjuk azokat.

Végezetül fel kívánjuk hívni a figyelmet elemzésünk egy másik korlátjára, amelynek a feloldása szintén jövőbeli kutatási irány lehet. Tanulmányunkban statisztikai hüvelykujjszabályként a 2 szórássterjedelmet alkalmaztuk az outlierok azonosítása céljából. Ez a „definíció” azonban önkényes választásunk eredménye, nem szükségképpen a legoptimálisabb választás. A bemutatott elemzések elvégezhetőek, és a kapott eredmények összevethetőek a kiugró értékek más lehetséges definícióival is.

Felhasznált irodalom

ALFARO, E. – GARCÍA, N. – GÁMEZ, M. – ELIZONDO, D. (2008): Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks. *Decision Support Systems*, vol. 45, pp. 110–122.

CAO, Y. (2012): MCELCCh: Financial distress prediction with classifier ensembles based on firm life cycle and Choquet integral. *Expert Systems with Applications*, vol. 39, pp. 7041–7049.

CHEN, N. – RIBERIO, B. – VIEIRA, A. – CHEN, A. (2013): Clustering and visualization of bankruptcy trajectory using self-organizing map. *Expert Systems with Applications*, vol. 40, pp. 385–393.

DU JARDIN, P. (2010): Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing*, vol. 73, pp. 2047–2060.

FEDOROVA, E. – GILENKO, E. – DOVZHENKO, S. (2013): Bankruptcy prediction for Russian companies: Application of combined classifiers. *Expert Systems with Applications*, vol. 40, pp. 7285–7293.

HORTA, I. M. – CAMANHO, A. S. (2013): Company failure prediction in the construction industry. *Expert Systems with Applications*, vol. 40, pp. 6253–6257.

KIM, M. J. – KANG, D. K. (2012): Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction. *Expert Systems with Applications*, vol. 39, pp. 9308–9314.

KRISTÓF, T. (2005): A csődelőrejelzés sokváltozós statisztikai módszerei és empirikus vizsgálata. *Statisztikai Szemle*, 83. évf. 9. sz. 841–863. o.

KRISTÓF, T. (2008): A csődelőrejelzés és a nem fizetési valószínűség módszertani kérdéseiről. *Közgazdasági Szemle*, LV. évf. 5. sz. 441–461. o.

MARQUÉS, A. I. – GARCÍA, V. – SÁNCHEZ, J. S. (2012a): Exploring the behaviour of base classifiers in credit scoring ensembles. *Expert Systems with Applications*, vol. 39, pp. 10244–10250.

MARQUÉS, A. I. – GARCÍA, V. – SÁNCHEZ, J. S. (2012b): Two-level classifier ensembles for credit risk assessment. *Expert Systems with Applications*, vol. 39, pp. 10916–10922.

NYITRAI, T. (2014): Validációs eljárások a csődelőrejelző modellek teljesítményének megítélésében. *Statisztikai Szemle*, 92. évf. 4. sz. 357–377. o.

ORESKI, S. – ORESKI, D. – ORESKI, G. (2012): Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment. *Expert Systems with Applications*, vol. 39, pp. 12605–12617.

PLATT, H. D. – PLATT, M. B. (1990): Development of a class of stable predictive variables: The case of bankruptcy prediction. *Journal of Business Finance and Accounting*, vol. 17, No. 1, pp. 31–44.

QUINLAN, J. R. (1993): *C4.5: Programs for Machine Learning*. San Mateo, California, Morgan Kaufmann.

SÁNCHEZ-LASHERAS, F. – DE ANDRÉS, J. – LORCA, P. – DE COS JUEZ, F. J. (2012): A hybrid device for the solution of sampling bias problems in the forecasting of firms' bankruptcy. *Systems with Applications*, vol. 39, pp. 7512–7523.

VIRÁG, M. – HAJDU, O. (1996): Pénzügyi mutatószámokon alapul csődmodell-számítások. *Bankszemle*, 40. évf. 4. sz. 42–53. o.

VIRÁG, M. – KRISTÓF, T. – FIÁTH, A. – VARSÁNYI, J. (2013): *Pénzügyi elemzés, csődelőrejelzés, vállalati válságkezelés*. Budapest, Kossuth Kiadó.

WANG, G. – MA, J. – YANG, S. (2014): An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, vol. 41, pp. 2353–2361.