

LAKATOS PÉTER–EGRI BOTOND–ASZALÓS PÉTER

VALÓSZÍNŰSÉGI HÁLÓ ALAPÚ BANKI CRM MODELL

A kereskedelmi bankokban egyeduralkodóvá vált az a szemlélet, amely szerint a tudatos ügyfélkapcsolat – különös tekintettel a dinamikusan bővülő hitelpiacra – szinte a kizárólagos módja lehet a versenyelőny megszerzésének, illetve megtartásának. A cikkben azt kívánjuk vázolni, hogy miként hozható létre egy, a banki igényekhez funkcionálisan jól illeszkedő ügyfélkapcsolati (customer relationship, CRM) modell.

Kereskedelmi banki környezetben különösen szerencsés ügyfélmodelleket létrehozni, ugyanis az, ahogyan az ügyfelek többsége a pénzügyeit kezeli, erősen magában hordozza az alapvető pénzügyi igényeikre, lehetőségeikre, tájékozottságukra, bankkal kapcsolatos attitűdjükre vonatkozó információkat. Az ügyfélmodellekre, illetve az azokat kiszolgáló informatikai rendszerekre a cikkben helyenként az általánosan elfogadott analitikus CRM kifejezéssel hivatkozunk.

Jelen tanulmányban érinteni kívánjuk az ügyfélkapcsolati modellek szinte minden lényegi alkotóelemét: kitérünk az **adatok** információtartalmának sűrítési lehetőségeire – miként juthatunk el az elemi tranzakciós adatoktól az ügyféligényekig; vázoljuk a rendszer magját alkotó **modell** üzletileg releváns funkcionalitását; végül elemezzük, miért szükséges a

modelleket egységes **informatikai infrastruktúrába** ágyazni.

ADATOK, FOGALMAK, SZEGMENSEK ÉS ÜGYFÉLIGÉNYEK

A kereskedelmi bankok analitikus CRM modelljeinek elsősorban a szakemberek által használt fogalmakat kell tartalmazniuk, amely fogalmak definiálásakor a heterogén és sokváltozós alapadatokat, illetve a szakértők tudását célszerű ötvözni. A fogalmi építőkövekből létrejövő modellnek pedig lehetőleg minél nagyobb leképezési szabadságot kell adnia.

Megközelítésünk két kulcselemet tartalmaz: egyrészt világosan elkülönítjük az adatok több rétegét (alapadatok – fogalmi szint – szegmens szint), másrészt a fogalmi szinten belüli és a fogalmi – szegmens

szint közötti sztochasztikus¹ kapcsolatok modellezését valószínűségi hálókkal valósítjuk meg; utóbbi előnyeit a cikk második részében vázoljuk.

A bankok igen nagyszámú **adat** rögzítenek az ügyfelekről, azok tranzakcióiról, pénzügyi portfólióikról, illetve ennek kezeléséről. Ez a több száz változó képezi a kereskedelmi banki analitikus CRM modell egyik építőkövét. A szóban forgó nagy elemszámú és heterogén változóhalmaz egyrészt meglehetősen ígéretes, ugyanis sok értékes információt hordoz az ügyfelekről, másrészt viszont óvatosságra int, ugyanis a közvetlenül ezekre épülő matematikai modellek sokszor széttöredezettek, nehezen interpretálhatóak és nincsenek összhangban azokkal a fogalmakkal, melyeket a bank szakemberei használnak. Szükségessé válik tehát ezen az adathalmazon olyan **dimenziócsökkentő transzformációk** elvégzése, melyek lehetőleg kevés információvesztéssel az ügyfelek pénzügyi lehetőségeit, bankkapcsolati sajátosságait, élethelyzetüket, illetve ahhoz kötött motivációit definiálják.

Ezzel a lépéssel eljutottunk a **fogalmi** szinthez. Az alapadatokból indultunk, melyekből egyszerű transzformációkkal, szabályokkal eljutottunk olyan fogalmakhoz, melyek már az ügyfelek pénzügyi interakcióit egy magasabb szinten írják le. A fenti fogalmi definíciós folyamat azon-

ban már nagymértékben támaszkodik a másik építőkövára: a bank szakértőinek tudására. A szakértők intenzív részvételével meghatározzuk a banki CRM szempontjából releváns fogalmakat, valamint azokat a transzformációkat, melyek leképezik az alapadatokat a megfelelő fogalmakba. A fogalmak használata ezáltal egyrészt tömörebb interpretációt tesz lehetővé, másrészt a szakértők számára megszokott környezetet teremt.

A fogalmak és alapadatok viszonyának szemléltetéséhez álljon itt egy példa az ügyfélkapcsolati csatornahasználathoz köthető ügyféljellemezőkről. Alapadatok szintjén rendelkezésre állnak információk arra vonatkozóan, hogy például az ügyfél rendelkezik-e elektronikus banki azonosítóval, használja-e és milyen gyakran, milyen mélységben; intézi-e telefonon a bankügyeit, igénybe veszi-e a mobilkészülékekről is elérhető biztonsági funkciókat, vagy pusztán bankfiókba jár, és egyszerűen felveszi a bankszámláján levő pénzügyösszeget anélkül, hogy kártyával fizetne, átutalásait kényelmesen, otthonról intézné. Az ügyfelek banki környezetben történő viselkedését az említett (több száz) változó meghatározza ugyan, de a kép értelmezéséhez a szakértő részéről jelentős emberi erőfeszítés szükséges. Ezzel szemben világos és könnyen értelmezhető képet kapunk abban az esetben, ha az alapadatokból kiindulva olyan fogalmakat használunk, mint például az ügyfél technológiai affinitása, készpénz- versus kártyapreferencia, bankkapcsolat szorossága, csatornadiverzitás. Pénzügyi adatok esetén például a diszkrecionális jövedelem, pénzügyi stabilitás, hitel-igénybevétel a

1 Determinisztikus változók közötti kapcsolat esetén egy kivételével minden más változó értékét rögzítve a nem rögzített változó értéke egyértelműen meghatározódik – például részecskék esetén a tömeg és az erőhatás ismeretében a gyorsulás. Sztochasztikus kapcsolat esetén a nem rögzített változó értékei valószínűség-eloszlást követnek – például jövedelem és életkor ismeretében a hiteltörlesztési fegyelem.

releváns fogalmak. Itt jegyezzük meg, hogy bár a fogalmak determinisztikusan épülnek fel az alapadatokból, a valószínűségi háló a fogalmak közötti kapcsolatokat már sztochasztikusan modellezi.

A szakértőkkel szoros együttműködésben definiált transzformációk alapján tehát létrejöttek a fogalmaink, melyek a **szegmens** szint alapját képezik. A szegmensek szintjét három lényegi elem különbözteti meg az előző, fogalmi rétegtől:

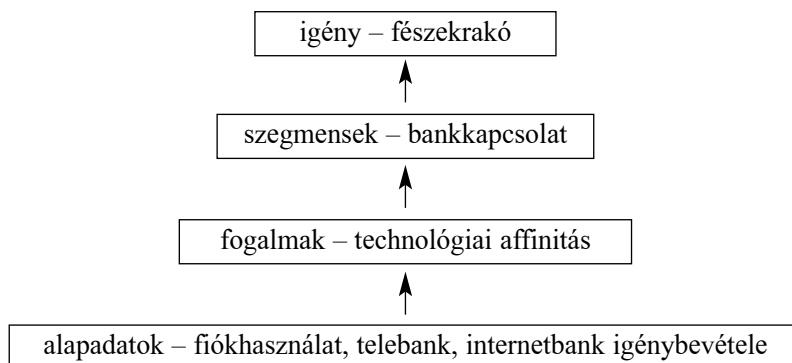
- a szegmensjelzők heterogén fogalmi csoportokra is épülhetnek;
- a fogalmak és a szegmensjelzők között sztochasztikusak a kapcsolatok, melyeket szintén az analitikus CRM modellt képező valószínűségi háló tartalmaz;
- a sztochasztikus kapcsolatokat a tanuló algoritmus szakértői segítség nélkül, önműködően építi fel, bár igény esetén a szakértők kényszereket definiálhatnak, az eredményként kapott modellt finomíthatják, szélsőséges esetben pedig kizárólag a tapasztalataikra alapozva, adatok nélkül is létrehozhatják.

Amint egyre magasabb szintre lépünk, úgy növekszik a változók által hordozott információ üzleti tartalma. A lojalitás, életvitel (hedonista avagy takarékos), életciklus (fészekrakó, nyugdíjas), pénzügyi helyzet, bankkapcsolat mind lehetséges, de korántsem egyedülálló szegmensjelzők, melyekből megtudhatjuk, hogy az egyes ügyfelek életpályájuknak mely szakaszában tartanak, valószínűségi becslést kapunk arra vonatkozóan, hogy mennyire lojálisak, illetve életvitelükhöz és pénzügyi helyzetükhöz igazíthatjuk ajánlatainkat [1].

A hierarchia legfelső szintjén helyezkedik el a motivációs (igény) szegmensjelző. Az ügyféligény, vagy motivációs szegmensjelzőt a bank szakemberei alakítják ki az alapján, hogy az egyes szegmensértékek különböző kombinációiból milyen mögöttes szükséglet meglétét valószínűsítik. Egy tipikus motivációs szegmensértékre példa a „fészekrakó” kategória, vagyis azon ügyfelek csoportja, akik meglehetősen lojálisak, hosszú távú be-

1. ábra

Változók hierarchikusan egymásra épülő rétegei



fektetéseket részesítenek előnyben, takarékosak, 20-as éveik végén, harmincas éveik elején járnak, családosak, otthonosan mozognak a különböző ügyfélkapcsolati csatornákon, gyakran vesznek igénybe különböző hiteltermékeket. A modell és a mögöttes matematikai algoritmus lehetővé teszi mindezen fogalmak kezelését, ezáltal a bank szakértői számára hatékony eszközt biztosítva az ügyfelek – valóban igényeikre és lehetőségeikre szabott – megszólításához.

Az igényekre (keresletre) épülő stratégia [2] a megszokott kínálati szemlélettől lényegesen eltérő megközelítést jelent. A többnyire piackutatási tevékenységre épülő kínálati megközelítés az alábbi folyamatot takarja: termékfejlesztés – árazás – reklám – eladás. Ezzel szemben az ügyfelek igényeinek mélyebb megértését a – tartós értéket hordozó konstrukció kialakítása követi – reklám – vásárlás.

A MODELLEK – VALÓSZÍNŰSÉGI HÁLÓK

A modellt alkotó építőelemek tisztázását követően térjünk most rá a modell szerkezetének, illetve ebből adódó előnyök megvilágítására.

A valós életben megfigyelt rendszerek nagy része sokváltozós és statisztikai modellekkel írható le. Ennek egyszerű oka az, hogy a gyakorlati életben vizsgált és érdekes rendszerek mindegyike több komponenst tartalmaz, valamint a komponensek között nem szigorú szabályok, hanem különböző ok-okozati csatornákon megvalósuló függőségek teremtenek viszonyt. A statisztikusságot tovább erősíti a

tény, hogy a rendszerek többsége – a gazdasági, szociális, politikai, de még a biológiai vagy kémiai rendszerek is – nyitott, azaz olyan hatásoknak van kitéve, amelyekkel nem tudunk, vagy nem akarunk számolni – egyszerűen nem tekintjük a rendszer részének. Ezek az korlátok azután a rendszerek leírására használt matematikai eszközök egy olyan osztályát jelölik ki, amelyek meg tudnak birkózni a fenti „lazasággal”: feltárják a változók közötti sztochasztikus összefüggéseket és ezen az úton létrehoznak egy modellt, amely már a rendszert leró adatok eredetétől függetlenül működtethető. Ezeket a modelleket grafikai modelleknek nevezzük.

A valószínűségi hálók [3, 4] szintén a grafikai modellek családjába tartoznak. A következőkben néhány gondolat erejéig bemutatjuk, hogy miért is jelent ez az algoritmuscsalád hatékony eszközt az ügyfélkapcsolati modellezéshez.

Komplex rendszerek modellezése.

Szerkezetükből adódóan kiválóan alkalmasak bonyolult rendszerek modellezésére és különböző rendszerekből adódó adatok integrálására. A kereskedelmi banki ügyfélkapcsolati modellek heterogén adatkörre épülnek: felhasználjuk a különböző alapszerekek tranzakciós adatait, kampányok eredményeit, a versenytársak aktivitását, az ügyfél visszajelzéseket – megteremtendő tehát ezek koherenciája.

Szakértői tudás integrálása. Az algoritmus az adatok alapján önműködően létrehozza a tárgyterületet leíró modellt, azonban a szakértők tapasztalataikra alapozva a későbbiekben ezen módosításokat eszközölhetnek. Szélsőséges esetben kizárólag a szakértők segítségével is lépésről lépésre

létrehozhatunk valószínűségi háló alapú intelligens asszisztenseket. Ez utóbbi szükségességéről már szoltunk az adatvagyon, illetve annak hasznosítását taglaló részben.

Maximális lekérdezési szabadság.

A klasszikus adatbányászati algoritmusok input-output rétegekbe rendezik a változókat, melyből adódóan a modellépítésnél kizárólag az outputként megadott változó(k)ra kérdezhetünk le, akkor, ha minden input változó értékét ismerjük. Ezzel szemben a valószínűségi hálók esetén nincsenek ilyen megkötések: bármely, modellbeli változóra lekérdezhetünk, akkor is, amikor csak néhány más változó értékét ismerjük. Klasszikus példa a fenti lehetőségre az ügyfél életciklus értékek [5] becslése olyan potenciális ügyfeleknél (prospect), akiről jellegükénél fogva nem ismerünk minden – az ügyfélérték modellezéshez egyébként felhasznált (például értékpapír-tranzakciók) – adatot, azonban a hiányos ismereteink birtokában is (például mérlegadatok, tevékenységi kör, foglalkoztatottak száma) lekérdezhetjük, hogy miként alakulna a potenciális ügyfél életciklus-

érték-eloszlása. Analóg módon a modell alapján nemcsak egy adott ügyfél lojalitására kaphatunk becslést, hanem fordított irányú kérdéseket is feltehetünk, például: az alacsony lojalitású, adott csatornahasználati szegmensbe tartozó ügyfeleimre nézve milyen a wallet-share.

Rejtett változók feltárása. Sokszor szembesülünk azzal az esettel, hogy általunk nem mérhető változók gyakorolnak hatást a rendszerünkre. Ezen rejtett változók bevetésével gyakran a tárgyterületet jobban leíró modellek hozhatók létre. A rejtett változókra a diszkrecionális jövedelem talán a legjobb példa.

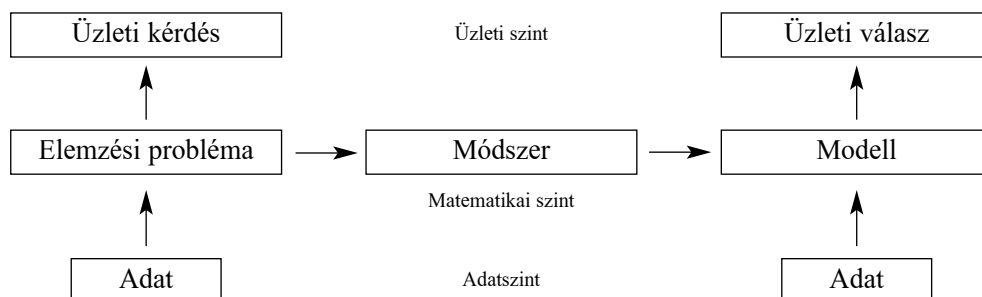
Modell alapú szegmentálás. A klasszikus távolság alapú szegmentálás nehézségei kiküszöbölő szegmentálási eljárás, melynek eredményeként pontosabb és könnyebben értelmezhető szegmensek jönnek létre.

AZ ELEMZÉSEK SZERKEZETE

Ebben a fejezetben az ügyfélkapcsolati modellezést folyamatában kívánjuk bemutatni. Ezúttal tehát nem elszigetelten az egyes al-

2. ábra

Elemzési folyamat általános sémája



kotóelemekre koncentrálunk, mint tettük azt az előző részekben az adatok és a matematikai modellcsalád esetében, hanem megkíséreljük megtalálni a szóban forgó elemek helyét az ügyfélanalizist igénylő folyamatban. A kérdést teljesen általánosan közelítjük azáltal, hogy ezúttal nem korlátozzuk vizsgálatunkat a kereskedelmi bank specifikus problémáira, hanem az elemzési folyamat általános szerkezetét kívánjuk bemutatni. Ez első olvasatra talán indokolatlanul messziről történő elrugaszkodásnak tűnhet, rávilágít azonban azokra a nehézségekre, melyek elkerülésére való törekvés motiválta a vizsgálatunk tárgyát képező ügyfélkapcsolati modelleket, illetve azt folyamatában támogató analitikus CRM-et.

Elemzési folyamat esetén – legyen szó a „Mely ügyfeleim bonyolítják a legnagyobb forgalmat?” kérdésről, vagy csavarva egyet rajta: „Mely ügyfeleim a legértékesebbek?” – alapvetően három réteget különböztethetünk meg:

- **üzleti szinten** fogalmazódnak meg a kérdések, illetve a szolgáltatott válaszok is itt hasznosulnak. Ez alatt helyezkedik el a
- **matematikai réteg**, melynek funkciója, hogy a maga nyelvezetét és eszköztárát használva megadja a választ az említett kérdésekre. A fentiek alapját az
- **adatszint** képezi, amelynek tartalmát a matematikai szinten kialakított modelleken átpréselve reményeink szerint eljutunk a válaszokhoz.

Lássuk most a folyamatot lépésről-lépésre: megszületik az üzleti kérdés, melyet az elemzők (vagy a szoftver) lefordítanak a matematika nyelvezetére. Kiválasztják a kérdés megválaszolásához releváns adato-

kat és a megfelelő módszert [6], majd kialakítják a probléma-specifikus modellt, melyet az adatokra alkalmazva a matematikai szinten eljutnak a válaszhoz, s ezt lefordítva az üzlet nyelvére a folyamat – legalábbis a visszacsatolásig – bezáru.

A kérdés – módszer – adat triumvirátusára álljon itt néhány környezetéből kiragadott és végletekig leegyszerűsített példa: ügyféllojalitás meghatározása – döntési fa – ügyféladatok; igények detektálása – független komponens analízis – tranzakciós adatok; szakértői rendszer – valószínűségi hálók – intuíció, tapasztalat; homogén célcsoport definiálása – Kohonen hálók – termékhasználat. Első olvasatra az étlap változatosnak tűnik, annak ellenére, hogy csak néhány alapfogást alkalmaztunk, ugyanis a feladatok túlnyomó többsége besorolható az osztályozás, klaszterképzés, interpretáció, előrejelzések valamelyikébe.²

A folyamat egyszerűnek tűnik, valójában tele van csapdahelyzetekkel. Vizsgáljuk meg ezeket közelebbről:

Interpretációs probléma. Az üzleti szinten és nyelvezetben megfogalmazott kérdéseket az elemzőnek interpretálnia kell, mégpedig oly módon, hogy a probléma matematikai megfogalmazásának az eredeti üzleti kérdéseket lehetőleg hűen kell tükröznie. A fordítási feladatokkal viszont sikerrel csak azok tudnak megbirkózni, akik mindkét nyelvet jól beszélnek.

2 Osztályozás esetén egy előre definiált kategóriába kívánjuk sorolni a megfigyeléseket – például lemorzsolódik vagy sem. Klaszterképzésnél valamely – szintén feltárandó – jellemzők mentén homogén csoportokat keresünk az adathalmazban – például csatornahasználat szerinti szegmensek. Interpretáció esetén többnyire a „Miért? Mivel jellemezhető?” kérdésre keressük a választ.

Információvesztés. Amennyiben a matematikai szinten az elemzési folyamat az üzletitől elszigetelten zajlik, könnyen előfordulhat, hogy az elemző olyan – üzleti szempontból egyébként hasznos – információt tár fel, melyet nem talál relevánsnak, s ez óhatatlanul elvész, elszivárog. Kiküszöbölésének egyetlen módja az, hogy a két réteg között szorosabbra kell fűzni az együttműködést és konzultációt.

Szakértői tudás integrálása. A szakértők az adatoknál többnyire nagyságrendekkel értékesebb információforrást képviselnek. Nem elhanyagolható szempont, hogy a tőlük származó információt sokkal könnyebb kinyerni, formalizálni azonban már nehezebb feladat. A szakértői tudásban megtestesülő információs vagyont szintén tudnunk kell integrálni az adott modellbe.

Időigényesség. A vázolt elemzési folyamat idő- és pénzigényes. A folyamatot szemléltető minden egyes nyíl akár többhetes időtartamot is takarhat, többszöri iterációkkal. Ennek megfelelően a jól standardizálható és gyakori üzleti kérdések esetén célszerű olyan megoldást találni, amely a fenti folyamatot minél inkább automatizálja.

Végtermék. A folyamat eredményeként sokféle végtermékhez juthatunk. Kaphatunk egymondatos választ; sorba rendezett ügyféllistát; százoldalas elemzést, vagy olyan alkalmazást, melynek végpontjai akár az ügyfélkapcsolati helyekre is kivezethetők.

Algoritmusok funkcionalitása. Végző soron a kérdést mégiscsak a modellt létrehozó algoritmusok pontossága, az adott problémára való megfelelése dönti el.

KIÚTKERESÉS

A veszélyforrások tisztázását követően belevághatunk a megoldás megformálásába: egységes rendszerre volna tehát szükségünk, melynek nem jelent gondot az adatintegráció, beágyazva tartalmazza a szükséges algoritmikus eszköztárat, közvetlenül az üzleti felhasználóhoz szóló front-end felülettel rendelkezik. Sikerült tehát kiiktatni az elemzőket, az üzlet maga dönt az információ relevanciájáról, gyorsan működik, szépen csiszolt modellt tartalmaz, az eredményeket adattárházban rögzíti, vagy akár a honlapunkon bóklászó ügyfelekhez vezeti ki. Megszületett tehát az analitikus CRM. Az igazsághoz hozzátartozik, hogy mindez bizonyos feltételek teljesüléséhez kötött: egyrészt feltételezi a rendszerhez intézett kérdések standardizálhatóságát, másrészt a kezdeti modell létrehozásához nem lehet megspórolni az egyszeri, de fáradságos adatbányászati elemzést igénylő munkaszakaszt, majd pedig a karbantartási munkálatokat. Vagyis az adatbányászati modellépítés folytatásaként létrehoztunk egy erre épülő rendszert is.

A feladatok azonban gyakran nem szoríthatók a fenti keretek közé, hiszen lehetnek egyediek, rosszul strukturáltak, különösen elemzőt próbálók, és – ami a leggyakoribb – expliciten nem körvonalazottak. Ez utóbbi eset képezi a klasszikus adatbányászati tevékenységet, vagyis prekoncepció nélkül alkalmazzuk az algoritmikus eszköztárat, azzal a céllal, hogy üzletileg releváns összefüggéseket találjunk.

KONKLÚZIÓK ÉS ÖSSZEFOGLALÁS

A modellek forrásoldalát tárgyalva kitérünk a szakértői tudás és az adatok egyenrangú szerepére. A modellek rendeltetése nem merül ki abban, hogy az adatok elemzésével új információt tárunk fel, hanem a szakértőkben rejlő tudást – ha úgy tetszik, a már meglévő információt – formalizált módon kinyerve, a modellek segítségével a teljes ügyfélkörre alkalmazhatjuk azt, mégpedig objektív módon.

Hangsúlyoztuk az adatok különböző üzleti információtartalommal bíró rétegein belül a fogalmi, illetve a motivációs réteg létrehozásának szükségességét. Az előbbi – elemi tranzakciós adatokból determinisztikus módon származtatva – tartalmazza a tárgyterület szakértőinek fogalomkészletét, melynek révén a modellek inputadatai szemléletesek, valamint eleve nagy magyarázóerővel bírnak. Ez utóbbi pedig megkísérli egyetlen változó értékeibe sűríteni ügyfeleink életciklusának állomásait, melyekre a banknak adekvát portfóliókat kell felkínálnia.

A valószínűségi hálók az ügyfélkapcsolati modellek létrehozásának kézenfekvő eszközei, mert mint láttuk, megfelelően rugalmas módon képesek kezelni a felmerülő kérdéseket.

Az elemzések szerkezetét taglaló részben arra kerestük a választ, hogy miért szükséges a modellezés alkotóelemeit egységes rendszerbe, alkalmazásba foglalni. Következtetésünk az volt, hogy standardizálható, gyakori kérdések esetén célszerű felparaméterezett analitikus CRM rendszert alkalmazni, amely magát alkotó matematikai algoritmus elegendően kifinomult ahhoz, hogy sikerrel vegye fel a küzdelmet az előtte tornyosuló feladatokkal. Nem elhanyagolható szempontként az üzleti felhasználóknak közvetlen hozzáférést biztosítottunk a modellekhez.

Láthattuk, hogy az ügyfélkapcsolati modellek felállításának nagyon komoly informatikai vonatkozásai vannak, felkészült szakembereket igényel a modellek matematikai kidolgozása, illetve ami talán a legnehezebb, a legtöbb időt és türelmet igénylő mozzanat: a sikerhez a szervezeten belül szemléletváltásnak kell következnie. Mindemellett tanulmányunkban nem tértünk ki a létrehozott eszközrendszer és modellek gyakorlati alkalmazására, holott a válaszok rendelkezésre állásától az ügyfelek megszólításáig a bank szakembereinek még sok tennivalójuk akad.

HIVATKOZÁSOK

- 1 Szép P., Hrotko T.: Jelenlét és a retail bankszolgáltatások más kihívásai. *Hitelintézet Szemle*, 2002. 1. szám.
- 2 Bank Administration Institute, The Cambridge Group Inc., Demand strategy in consumer financial services, 2001.
- 3 Heckerman, D.: A Tutorial on Learning with Bayesian Networks., *MSR-TR-95-06*, 1995.
- 4 Russell, S. J., Norvig, P.: Mesterséges intelligencia modern megközelítésben, pp. 522–545., *Panem-Prentice Hall*, 2000.
- 5 Lakatos P., Cseh Z. Churn and customer value. *Proceedings of CLUG Reading*, UK., 2000.
- 6 Cios, K. et al., Data Mining – Methods for knowledge discovery. *Kluwer*, 1998.