



MKKV ügyfelek adóminősítő modelljének fejlesztése

RapidMiner a Takarékbankban

Frindt Anna
Magyar Takarékszövetkezeti Bank Zrt.





A Takarékbank és a Takarékszövetkezetek/Bankok

1989 – Magyar Takarékszövetkezeti Bank Zrt. alapítása

Alapítók – Takarékszövetkezetek

Ma is a Takarékszövetkezetek/takarékszövetkezetekből átalakulással létrejött Bankok többségi tulajdonában

Kitűzött feladatok: a takarékszövetkezetek egységes piaci fellépésének megteremtése, versenypozíciójuk megerősítése, forrásaik hasznosítása, valamint tevékenységi korlátaik feloldása.

Takarékszövetkezeti szektor központi bankja.

Alapvető ügyfélkör: takarékszövetkezetek, kis- és középvállalkozások, önkormányzatok.



1993 – Takarékszövetkezeti Integráció megalakul

A Takarékbank, valamint – ma – több, mint 100 önálló takarékszövetkezet/bank együttműködése
Kiterjedt fiókhálózat (kb. 1700) az ország területén.

Legjelentősebb ügyfelei a mikro-, kis-, és középvállalkozások és a lakosság.

Jellemző a vidéki ügyfelek nagy száma, jelentős hányaduk az agrárszférába tartozik.



2008 – TakarékJpont együttműködés elindul

Az együttműködésben résztvevő önálló hitelintézetek egységes arculattal és lakossági szolgáltatásokkal állnak országsszerte az ügyfelek rendelkezésére.

Jelenleg 24 takarékszövetkezet/bank a tagja.





Hitelezési kockázat, adósmínősítő modellek

☉ A **hitelezési kockázat** a bank tőkéjéhez viszonyítva a legjelentősebb kockázatforrás.

☉ **A banknak meg kell vizsgálnia**

☉ **Új ügyfelek esetében:**

- Kinek adjon hitelt?
- Mekkora összeget hitelezzen?
- Mekkora futamidőre?
- Milyen biztosítékokkal?
- Milyen kockázati felárral folyósítson hitelt?

☉ **Meglevő ügyfelek esetében:**

- Megváltozott-e az adós kockázati megítélése?

☉ Ezért az ügyfeleket **minősíteni kell** több szempont alapján:

- Objektív, számszerűsíthető elemek (pl.: pénzügyi mutatók, alapítás óta eltelt idő)
- Szubjektív, nem számszerűsíthető elemek (pl.: vállalati vezetés színvonala, banki kapcsolat minősége)

☉ **Adósmínősítő modelleket** az objektív elemek alapján statisztikai módszerekkel lehet fejleszteni.

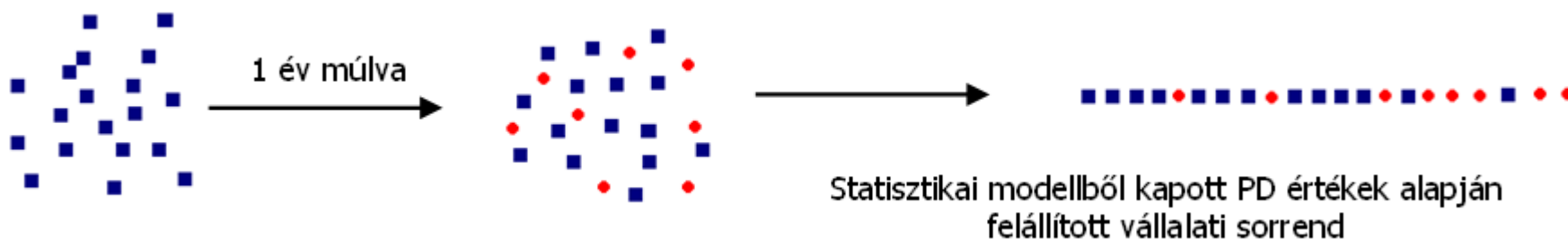
☉ Az ügyfelek szubjektív minősítését hitelezési tapasztalattal rendelkező kockázatelemző végzi.





Adósminősítő modellek elméleti háttere

- ☉ **PD (probability of default):** annak a valószínűsége, hogy egy cég a következő egy évben bedől. A modellel ezt számítjuk ki.
- ☉ **A felhasznált adatok:** „kockázati tényezők”, azaz magyarázó változók és a bedőlés ténye.
- ☉ Egy modell **pontszámot (score) ad** a modellezési eljárásban felhasznált cégeknek. A legjobb modellt egy statisztikai program választja ki úgy, hogy a bedőlteket ne keveredjenek a nem bedőltekkal (lásd ábra). Ezt tesztmintán is ellenőrizzük.
- ☉ A kész modellel egy **országos cégmintát pontozunk**, és a cégeket a pontszám alapján sorba rendezzük. Ha ezután egy céget minősíteni akarunk, akkor a modellel pontozzuk, és megnézzük, hogy hová esik az országos rangsorban. A hasonló pontszámot kapott cégek körében mért tényleges múltbeli bedőlési arány lesz a minősített cég PD-je.





Adóminősítő modellekhez használt adatállományok MKKV ügyfelek esetén

🌐 **Bedőlés adatok, késedelmes fizetést mutató adatok** (Igen/Nem mutató):

- Bukó vállalatok
- Nembukó vállalatok
- Szürke vállalatok

🌐 **Magyarázó változók, a bedőléssel összefüggésbe hozható kockázati tényezők:**

- Vállalkozás törzsadatai:
Székhely
Alapítás dátuma
Főtevékenység stb.
- Vállalkozás pénzügyi adatai az elmúlt 2 évben:
Mérleg
Eredmény-kimutatás
- Vállalkozás egyéb céginformációs adatai:
Bankváltás gyakorisága
ISO minősített-e
EU-s pályázatot nyert-e stb.





Adóminősítő modellek fejlesztése során felmerülő problémák

- ⊗ **Adattisztítási lépések**
 - Rengeteg hibás adat miatt szűrő funkciók szükségesek
- ⊗ **Modellezési minta összeállítása**
 - Rétegzett mintavételi eljárás szükséges
- ⊗ **Független tesztelési minta elkülönítése**
 - Modellezési minta kettéválasztása
- ⊗ **Magyarázó változók diszkretizációja**
 - Folytonos változók kategória-értékké alakítása
- ⊗ **Hiányzó adatok pótlása**
 - További adatvesztés elkerülése, 0-val való osztás kezelése
- ⊗ **Statisztikai modellezési eljárás**
 - Logisztikus regresszió
- ⊗ **Modell teljesítményének tesztelése**
 - Országos mintán való ellenőrzés





Korábbi modellezési eszközök

☛ Korábban használt eszközök:

- Statistica
- Excel
- R



☛ A korábbi eszközök használata során felmerült problémák:

- Statistica: logisztikus regresszió nem működik megfelelően, gyakori lefagyás
- Excel: nem alkalmas nagy adatállományokkal való bányásra

☛ Szerteágazó környezet miatt magas modellkockázat

☛ Minden elemből azt a funkciót használtuk, amire a legalkalmasabb volt:

- Statistica: adatállományok összefésülése, rendezése (adatkezelési munka)
- Excel: egyszerű kalkulációk
- R: statisztikai eljárások (modellezési munka)





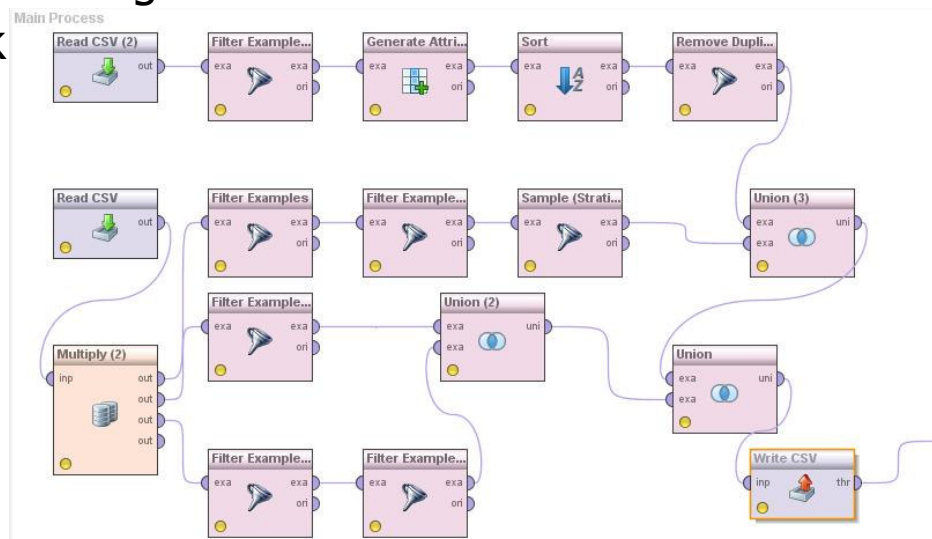
Miért pont a RapidMiner? Előnyök és hátrányok

Előnyök:

- R integrációja megoldott
- Weka fejlesztéseit tartalmazza
- Community Edition – díjmentes használat
- Felhasználói fórumok
- Felhasználóbarát kezelő felület
- Modulokból felépülő folyamatokra megjelenítése
- Automatikus ellenőrző funkció és javítási lehetőségek
- Logisztikus regresszió megfelelően működik (stepwise módszer is)
- Loggolás

Hátrányok:

- Lassú tanulási folyamat
- Kevés dokumentáció
- Update nehézkes a banki környezetben





Köszönöm a figyelmet !

