

## **ÉRVÉNYESSÉG ÉS KORLÁTOK AZ ALGORITMIKUS DÖNTÉSHOZATALBAN**

*Hámori Gábor*

Manapság nem meglepő egy internetes könyvruházban keresgélve, hogy az oldal automatikusan felkínál számunkra bizonyos könyveket, amelyek jó eséllyel számíthatnak az érdeklődésünkre. Telefonszolgáltatónktól sms-t kapunk, amelyben célirányosan javasolnak új terméket számunkra. Banki hitelért folyamodva, kérelmünket ma már pár perc alatt automatikusan elbírálja a hitelintézet. A felsorolt példákban az a közös, hogy minden esetben algoritmikus prediktív módszerek segítségével történtek előrejelzések várható viselkedésünkre. De vajon akkor is sikeresen alkalmazhatók-e a matematikai-statisztikai módszerek, ha például a bordeaux-i borok jövőbeli áralakulására vagy egy házasság tartósságára szeretnénk prognózist felállítani? Vagy ezekben az esetekben inkább bízunk meg az adott területen tapasztalattal rendelkező szakértőkben? Az alábbi tanulmány segítségével a témában megjelent kutatásokra támaszkodva, röviden áttekintjük, melyek azok a tényezők, amelyek meghatározzák, hogy egy adott döntési-előrejelzési szituációban algoritmikus eszközökre vagy szakértői becslésekre támaszkodjunk-e inkább.

*JEL-kódok:* C10, C30, D70, D81

*Kulcsszavak:* döntésemélet, kognitív torzítás, predikció, statisztikai előrejelzés, szakértői becslés

### **A DÖNTÉSEINKET BEFOLYÁSOLÓ GONDOLKODÁSI HIBÁK**

Hogy a döntésekben betöltött szerepük alapján az algoritmikus döntési módszereket megfelelően tudjuk értékelni, először a teljesség igénye nélkül vegyük sorra azokat a tipikus gondolkodási hibákat, amelyeket jellemzően nap mint nap, öntudatlanul elkövetünk, és hatásuk alól a szakértői értékítéletek sem mentesek, ezáltal befolyásolják a kapcsolódó döntések és előrejelzések minőségét.

Talán az összes gondolkodási hiba közül a legjelentősebb hatású az úgynevezett *megerősítési torzítás*, ami azt jelenti, hogy döntéseinket megelőzően a rendelkezésre álló információkat hajlamosak vagyunk úgy értékelni, hogy lehetőség szerint ne kerüljenek összeütközésbe a problémával kapcsolatos szakmai meggyőződésünkkel, nézeteinkkel. Ennek következtében jellemzően azokat az információkat, amelyek nincsenek összhangban világnézetünkkel, kiszűrjük, vagy lényegesen kisebb súllyal vesszük figyelembe, mint az valójában indokolt lenne.

Erre utal *Warren Buffett* elhíresült mondása is: „Az emberek ahhoz értenek a legjobban, hogy az új információkat úgy szűrjék meg, hogy meglévő nézeteik érintetlenek maradjanak.” A legnagyobb probléma az, hogy a megerősítési torzítás általában nem tudatosul, az elméleteinknek ellentmondó információkat többnyire „meg sem halljuk”.

Gyakran vezet hibás döntésekhez a *problémát leíró valószínűségek figyelmen kívül hagyása* és ezen belül is a *zéró kockázat iránti elfogultság* (Slovic, 2000). Például amikor a lottófogadók és más szerencsejátékosok beszállnak a játékba, minden bizonnyal figyelmen kívül hagyják a játék valós valószínűségi viszonyait, ezáltal válik különlegesen jövedelmezővé a szerencsejáték-ipar. A zéró kockázat iránti elfogultság megértéséhez tegyük fel, hogy két projekt közül választhatunk (Dobelli, 2013). Az egyik esetben a környezetszennyezés okozta halálesetek számát 5 százalékról 2 százalékra csökkenthetjük. A másik azonos feltételrendszerű (költségek, időhorizont, stb.) projekt segítségével az 1 százalékos halálozási kockázat teljesen megszüntethető. A legtöbb ember a második alternatívát részesíti előnyben, dacára annak, hogy az első esetben a javulás 3 százalékpont, azaz háromszorosa a zéró kockázatot igényő beruházásnak.

A valószínűségek figyelmen kívül hagyásának egy egyszerűbb, de annál gyakoribb esete az *alapgyakoriság figyelmen kívül hagyása*. A következő példa segítségével könnyen érthetővé válik, miről is van szó tulajdonképpen. Tegyük fel, hogy Péter vékony, szemüveges férfi és szereti Bachot. Mi a valószínűbb? Péter *a)* taxisofőr; vagy *b)* irodalomprofesszor Budapesten. A legtöbben a második lehetőségre tippelnek, tévesen. Magyarországon nagyságrenddel több taxisofőr él, mint budapesti irodalomprofesszor. Ezért lényegesen valószínűbb, hogy Péter taxisofőr, még ha szereti is Bachot. A második válaszra tippelők figyelmen kívül hagyták az alapgyakoriságot.

Elsősorban bonyolult rendszerek hosszú távú szakértői előrejelzései esetében figyelhető meg az úgynevezett *prognózisillúzió* (Dobelli, 2013). *Philip Tetlock*, a Berkeley professzora 284 szakértő több mint nyolcvanezer előrejelzését értékelt (Tetlock, 2006.) A vizsgálat kimutatta, hogy az előrejelzések alig voltak valamivel jobbak, mintha véletlenszerűen hozták volna létre a jövőre vonatkozó prognózisokat. A kutatás meglepő eredménye volt az, hogy leginkább a kiemelt médiafigyelmet élvező szakértők jóslatai bizonyultak megbízhatatlannak. *John Kenneth Galbraith* harvardi közgazdász kissé szarkasztikusan a következőképpen sommazta a jelenséget: „A jövőbelátóknak két típusuk van: azok, akik nem tudnak semmit és azok, akik nem tudják, hogy nem tudnak semmit.”

És végezetül említsük meg a *túlzott magabiztosságot*, ami azt jelenti, hogy jellemzően túlbecsüljük saját tudásunkat. Meglepő módon a túlzott magabiztosság a szakértőkre, és ezen belül is a férfiakra átlagon felül jellemző. Ez utóbbinak feltehetően evolúciós okai is lehetnek (Baumeister, 2001). Nem szorosan a köz-

gazdaságtudomány körébe tartozó példa erre egy Franciaországban elvégzett kutatás, amelyben a francia férfiak 84 százaléka nyilatkozott úgy, hogy átlagon felüli szerető (*Dobelli, 2013*). A kérdésben közel normális eloszlást feltételezve, a valóságban 50 százalék körül lenne ez a szám, amennyiben nem létezne a túlzott magabiztosság effektus. Vélhetően ehhez hasonló eredményt kapnánk, ha a szakértőket kérnénk meg arra, hogy pozicionálják önmagukat és tudásukat az azonos szakterületen működő kollégáik alkotta eloszlásban.

Egyéb tényezők is befolyásolják gondolkodásunkat, ezáltal döntéseink minőségét. Ugyanazt a döntési helyzetet két különböző időpontban kedélyállapotunktól, napszaktól és számos egyéb tényezőtől (pl. fáradtságunktól) függően nem ugyanúgy értékeljük, azaz értékítéletünk időben *inkoherens*. Az inkoherens döntések ugyanúgy jellemzik a hétköznapi helyzeteket, mint a szakértők által támogatott döntési szituációkat. Erre jó példa az a kutatás (*Hoffman–Slovic–Rorer, 1968*), amelynek során tapasztalt radiológusokat kértek meg arra, hogy mellkasröntgen-felvételeket értékeljenek a „normális” és „abnormális” kategóriák mentén két különböző időpontban; a kísérletben résztvevők nem tudták, hogy másodszorra ugyanazokat a felvételeket látják. Az esetek 20 százalékában az értékelések önelmentmondóak voltak. Hasonló mértékű következetlenséget figyeltek meg abban a kutatásban is, amelynek során 101 könyvvizsgálót kértek fel arra, hogy értékeljék a vállalati belső ellenőrzések megbízhatóságát (*Brown, 1983*).

## MIKOR HIHETÜNK A SZAKÉRTŐKNEK?

Miután röviden áttekintettük a jellemző gondolkodási hibákat, amelyek a szakértők előrejelzését és ítéletalkotását is jelentősen torzíthatják, felmerül a kérdés: mikor és milyen körülmények között bízhatunk egy szakértői véleményében, intuícióiban? A kérdés megválaszolásához vegyük figyelembe, hogy a szakértelem tulajdonképpen hosszas tanulás során megszerzett, számos releváns készség összessége. Ez teljesen nyilvánvaló például egy profi sakkozó esetében. A magas szintű sakkismeretek megszerzéséhez éveken keresztül tartó, napi többórás gyakorlásra volt szükség. Nincs ez másképpen egyéb szakterületeken sem. A szakértői értékítéletek érvényességének két alapfeltételét adja meg *Kahneman* (*Kahneman, 2013*):

- megfelelően szabályos és kiszámítható környezet;
- hosszas gyakorlás, amely lehetőséget biztosít a megfelelő készségek elsajátítására és a szabályszerűségek megtanulására.

Az első feltétel a sakkjáték esetében nyilvánvaló, de hasonlóan szabályos környezetben dolgoznak a mérnökök és a klinikai orvosok is. Ezzel szemben a hosszú távú előrejelzésekkel dolgozó politológusok, értékpapír-elemzők alapvető-

en zajos, nagyon alacsony rendezettségű környezetben működnek. A második feltétel teljesülése nagyban függ a tevékenység visszacsatolásának természetétől. *Kahneman* ennek megvilágítására a szakorvosok példáját hozza: „A szakorvosok közül az aneszteziológusoknak van része a legjobb visszacsatolásban, mert tevékenységük hatásait nagy valószínűséggel azonnal láthatják. Ezzel szemben a radiológusok kevés információt nyernek a felállított diagnózisok pontosságáról vagy a fel nem ismert betegségekről. Ezért az aneszteziológusok könnyebben ki tudják alakítani a megfelelő intuitív készségeket. Ha egy aneszteziológus azt mondja: »Úgy érzem, valami nincsen rendben«, a műtőben mindenkinek fel kell készülnie az azonnali beavatkozásra.” Ennek alapján az is könnyen érthetővé válik, hogy a szakértői előrejelzések jellemzően miért jobbak rövid időhorizonton, mint hosszabb távon. A hosszabb távú előrejelzéseket egyszerűen nincsen módunkban elégszer gyakorolni.

## JÖHET A STATISZTIKA!

Az emberi gondolkodás és a szakértői intuíciók néhány, a döntési helyzetekben meghatározó jellemzőjének áttekintése után vizsgáljuk meg, hogy milyen eredmények érhetők el matematikai-statisztikai eszközök használatával.

A számítástechnika területén megfigyelhető technológiai forradalomnak az elmúlt évtizedekben megfigyelhető, robbanásszerű fejlődése lehetővé tette, hogy másodpercek alatt több millió műveletet hajthassunk végre személyi számítógépek segítségével. Korábban a bonyolult matematikai-statisztikai algoritmusok (továbbiakban: algoritmusok) alkalmazása a gyakorlatban inkább csak az akadémiai kutatások körében volt tetten érhető, végrehajtásuk jelentős időigénye és a korlátozott számítástechnikai kapacitás miatt. Mára azonban a helyzet gyökeresen megváltozott. Külön iparág jött létre azzal a céllal, hogy az algoritmusokat a felhasználók részére elérhetővé tegye. Rendre elérhetővé váltak a különböző célszoftverek, amelyek segítségével a korábban akár hetekig tartó elemzéseket pár perc alatt végrehajthatjuk. Hála a nyílt forráskódú szoftvereknek, ma már az algoritmusok jelentős része bárki számára hozzáférhető.

A számítási lehetőségek gyors ütemű bővülése és a könnyű hozzáférés természetesen további lökést adott az alapkutatásoknak, ezen belül is az algoritmusfejlesztésnek. Szakmai tudományos műhelyekben rendszeresen lát napvilágot egy-egy új eljárás vagy algoritmus, amelyek közül a legéletképesebbek a gyakorlatban is hamar meghonosodnak.

Az alkalmazás területén elsőként a vállalati/üzleti szféra ismerte fel az új technológiában rejlő lehetőségeket. Az olyan iparágak, ahol jellemző a tömegigényeket kielégítő termelés, illetve szolgáltatás, az üzleti folyamatok statisztikai elemzése

és modellezése magától értetődő. A tömegszerű döntéseket igénylő helyzetekben a korszerű nagyvállalatok korán felismerték, hogy algoritmusok eredményeképpen előálló, prediktív döntési szabályok/formulák segítségével gyorsan, és ami talán még ennél is fontosabb, olcsón tudnak dönteni.

Azokban a szituációkban, amikor a vizsgált probléma komplexitása, „mérete” jelentős, a döntési helyzet összetettsége következtében az algoritmusok olyan összefüggések feltárására is képesek, amelyeket az emberi elme már nem tud megragadni. Ez különösen igaz akkor, ha például az előrejelzendő mennyiség alakulására számos faktor hat, de egyenként viszonylag kicsi súllyal befolyásolják a végeredmény bekövetkezését. Gondoljunk például egy félmillió ügyféllel és ennek megfelelően rengeteg leíró adattal rendelkező hitelintézetre, ahol arra keresik a választ: mitől függött és hogyan, hogy valaki visszafizette-e a kapott hitelt, vagy sem? Vagy lehetséges-e valószínűségi becslést adni egy új ügyfél esetében a nemfizetés valószínűségére, és ennek alapján dönteni, hogy ki kapjon hitelt?

Megfelelő statisztikai modellek segítségével jellemzően feltárhatók mindazok a tényezők, amelyek hatást gyakorolnak a fizetési hajlandóságra, valamint alkalmas algoritmusok segítségével előállítható olyan formula, amelynek a segítségével meg lehet becsülni egy konkrét ügyfél esetében a nemfizetés valószínűségét. A feltárt összefüggések alapján olyan üzleti döntési szabályok hozhatók létre, amelyek egy célfüggvény mentén optimális folyamatot eredményeznek, és azok az alkalmazás során automatikusan végrehajtásra kerülnek.<sup>1</sup>

A fenti döntési helyzet összetettsége okán talán nem meglepő, hogy ezeken a területeken az emberi (szakértői) döntés/becslés rosszabbul teljesít, mint az algoritmusalapú szabály. A meglepő azonban az, hogy sok esetben ugyanez figyelhető meg „kis mintás” esetben is, azaz ott, ahol a probléma jóval kevesebb adattal írható le. *Paul Meehl* 1954-ben megjelent *Clinical vs. Statistical Prediction: A Theoretical Analysis and a Review of the Evidence* című munkájában 20 olyan kutatás eredményeit foglalta össze, amelyek azt vizsgálták, hogy a képzett szakemberek szubjektív értékítéletén alapuló klinikai előrejelzések pontosabbak-e, mint egy adott szabályból származtatható, statisztikai előrejelzés. A kutatás eredménye szerint a statisztikai előrejelzések jellemzően sokkal jobban teljesítettek, mint a szakértők általi becslések. Még meglepőbb a 2002-ben közgazdasági Nobel-díjjal kitüntetett *Daniel Kahneman* eredménye a témában, aki 2011-ben megjelent *Gyors és lassú gondolkodás* című művében a következőket írja: „A klinikai és statisztikai előrejelzések összehasonlításáról beszámoló kutatások száma mintegy 200-ra nőtt, de az

<sup>1</sup> A fenti két kérdéshez kapcsolódóan jó példa a nagybankok által használt automatikus hitellel bírálati, ún. scoringrendszerek alkalmazása. Ezek a rendszerek a megadott adatok alapján teljesen automatikusan, az emberi tényező teljes kizárásával, másodpercek alatt elvégzik a hitelkérelmező ügyfél nemfizetési valószínűség-becslését, és ennek alapján tesznek javaslatot a szerződés megkötésére, illetve magas nemfizetési valószínűség esetén a kérelem elutasítására.

algoritmusok és az emberi előrejelzések közötti verseny eredményei nem változtak. A kutatások közel 60%-ban az algoritmusokat jelentősen pontosabbnak találták. A többi összehasonlítás döntetlen eredményt hozott a pontosság szempontjából, de a holtverseny valójában a statisztikai szabályok győzelmét jelentette, mivel ezek általában sokkal olcsóbbak, mint a szakértői ítéletek. A kutatások egyetlen meggyőző kivételről sem számoltak be.<sup>2</sup> További váratlan eredménye a kutatásoknak, hogy a statisztikai predikciók fölénye még a magyarázó faktorok egyszerű lineáris kombinációjaként előálló szabályok esetén is megfigyelhető.

A statisztikai előrejelzések fölénye olyan területeken is tetten érhető, ahol a legkevésbé várnánk. A princetoni közgazdász és borkedvelő *Orley Ashenfelter* (2007) statisztikai előrejelzése a bordeaux-i borok árának alakulására felülmúlta a legelismertebb borszakértők prognózisait is. *Ashenfelter* ehhez a borok készítésének évében megismerhető információkat használta. Érdekességképpen, a végleges előrejelző regressziós modell három időjárási jellemzőt tartalmazott: a nyári természetes időszak átlaghőmérsékletét, a csapadék mennyiségét szüretkor, illetve az összes csapadék mennyiségét az elmúlt tél során. Pusztán ezzel a három magyarázó változóval *Ashenfelter* a bordeaux-i borok jövőbeli ár alakulásának varianciáját  $R^2 = 82\%$ -ban meg tudta magyarázni.

Hogy *Ashenfelter* megközelítése mennyire nem elszigetelt jelenség, azt az alábbi, csak felsorolásszerűen említett néhány kutatás mutatja az elmúlt évtizedekből, amelyek egy adott előrejelzési kérdésben a statisztikai predikciókat vetik össze a szakértői előrejelzésekkel.

- *Howard* és *Dawes* (1976) jól és a lehető legegyszerűbb módon jelezte előre a házasságok várható stabilitását a következő képlettel:  $P = \text{szeretkezések gyakorisága} - \text{veszekedések gyakorisága}$ . A képlet megbízhatóságát később *Edwards* (1977) és *Thornton* (1979) ellenőrizte.
- *Wittman* (1941) elektrosokk-terápia sikerességének előrejelzésére alkalmazott statisztikai modellt, amelynek előrejelzési képessége meghaladta az eljárást alkalmazó pszichológusok prognózisait.
- *Carroll* (1988) statisztikai előrejelző módszerekkel sikeresebben jelezte előre az elítéltek visszaesését, mint a tapasztalt kriminológus szakértők.
- Az egyes diákok esetében a felvételi pillanatában készített statisztikai modellek jobban jelezték előre az egyetemi tanulmányok terén várható előmenetelt – legyen szó orvostudásról (*DeVaul* et al., 1957) vagy jogászképzésről (*Swets–Dawes–Monahan*, 2000), mint a felvételiztető tanárok.
- Banki hitelezés területén a nemfizetés előrejelzésében a statisztikai modellek elsőbbségét a hitelezési szakértőkkel szemben *Stillwell* (1983) és *Libby* (1976) tanulmánya is megerősíti.

2 KAHNEMAN (2013), 258. o.

- Több tanulmány is rámutatott arra, hogy a korai csecsemőhalál (bölcsőhalál) bekövetkezése jobban előrejelezhető statisztikai modellek segítségével, mint a szakértők által (Lowry, 1975; Carpenter, 1977; Golding, 1985).
- Statisztikai modellek segítségével jobban lehetett előre jelezni egy adott egyén esetében a jövőbeli bűnelkövetés valószínűségét, mint azt a témában jártas pszichiáterek tették (Faust–Ziskin, 1988).

A statisztikai előrejelzések alapjául szolgáló modellek a jellemzően magasabb teljesítményen túl arra is alkalmasak, hogy feltárják az előrejelzés magyarázó faktorait, azok összefüggéseit egymással és az előre jelezni kívánt mennyiséggel. Ezáltal nemcsak jól működő „kristálygömbként” hasznosíthatók az előrejelzés során, hanem általuk tudásra, a modellezendő jelenség belső mechanizmusának és az összefüggések rendszerének megértésére is szert tehetünk. Leegyszerűsítve azt is mondhatnánk, hogy egy megfelelő adatbázison egy jó statisztikai modell pár perc alatt képes „megtanulni” mindazt, amihez a terület szakértőjének esetleg több évtizedre volt szüksége

## A STATISZTIKAI ELŐREJELZÉSEK ÉRVÉNYESSÉGE

Az eddigiekből talán arra következtethetünk, hogy a legtöbb döntési és előrejelzési szituációban teljesen mellőzhető az emberi bölcsesség, a szakértői tudás. „Szerencsére” a statisztikai alapú algoritmikus döntési és előrejelzési rendszerek sem mentesek bizonyos hátrányoktól. Az egyik ilyen a Paul Meehl által „törött láb” jelenségnek (broken leg phenomenon) nevezett helyzet. Meehl gondolat kísérletében feltette, hogy rendelkezünk egy statisztikai algoritmussal, amely a korábbi tapasztalatok alapján nagy biztonsággal képes előre jelezni, hogy egy bizonyos professzor szerda este moziba fog menni. Az algoritmus remekül működik mindaddig, míg egy keddi napon a professzor váratlanul el nem törí a lábát, és így szerdán nem tud moziba menni. Tehát az algoritmusok rosszul teljesítenek olyan helyzetekben, amikor bekövetkezik egy korábban meg nem figyelt, alacsony valószínűségű, ritka esemény, amelynek a kimenetre gyakorolt hatása jelentős.<sup>3</sup>

A másik jellemző probléma az algoritmusfejlesztés kiindulását jelentő adatszerkezettel kapcsolatos. Az előrejelző modellek csak az adatok által reprezentált világban képesek optimális eredményeket szolgáltatni. Ha valamilyen oknál fogva az adatbázisból kimaradnak lényeges, a prediktálni kívánt mennyiség alakulását erősen befolyásoló változók, az előrejelzés minősége romlani fog. Ilyen helyzetre példa, amikor egy hitelintézet bizonyos típusú adatokat (pl. vallás, rassz stb.) az ügyfelektől az adatvédelmi szabályoknak megfelelően nem gyűjthet.

3 Gazdasági rövid távú előrejelzésekben ilyen atipikus esemény a trendforduló vagy a válsághelyzet.

A gyakorlati alkalmazásnál további fontos szempont az adatok megbízhatóságának, minőségének a kérdése, amit elsősorban az adatrögzítési és adatkezelési folyamatok határoznak meg. A felhasznált adatok minősége, megbízhatósága meghatározó a modell előrejelzési képessége szempontjából. Ha nem valós, hibás adatokkal szennyezett adatszerkezeteket használunk a statisztikai modellezés során, az eredmény is korlátozottan lesz érvényes. Ezt a jelenséget fejezi ki talán kicsit túlzóan az úgynevezett GIGO-elv (Garbage In, Garbage Out).<sup>4</sup>

A statisztikai modellek alkalmazásának és elterjedésének egyfajta természetes, az emberi pszichikumból származó korlátja a velük szemben nemegyszer tapasztalható idegenkedés, ellenérzés. Jó példa erre a már korábban említett *Ashenfelter* bordeaux-i borok árát előrejelző képletének fogadtatása. A francia borbarátok reakciója a *New York Times* szerint az „agresszívtól a hisztérikusig” terjedt. *Kahneman* szerint az algoritmusok iránti előítélet tovább nő, amint az ezekre alapozott döntéseknek jelentős következményei vannak. A legtöbb ember számára ugyanis nem közömbös, hogy a valós helyzeteket jellemző döntési hibák forrása egy algoritmus vagy egy szakember. Egy orvosi műhiba esetén megrendítőbb, ha egy beteg elhalálása egy képlet alkalmazása miatt következik be, mintha egy orvos rossz döntése vezetett volna a sajnálatos eseményhez.

## ALGORITMUS VAGY SZAKÉRTŐK?

Végezetül próbáljuk meg áttekinteni, hogy egy adott döntési helyzetben statisztikai eszközökhöz folyamodjunk-e, vagy inkább a szakértők tanácsára támaszkodjunk.

*Zajos, alacsony strukturáltságú előrejelzési környezetekben*, mint például az értékpapír-árfolyamok világa, a kutatások tanulsága szerint sem a matematikai-statisztikai algoritmusok, sem a szakértői előrejelzések nem teljesítenek megfelelően, függetlenül az időhorizonttól. Ebben az esetben a legnagyobb problémát a megfigyelt rendszer szabályszerűségeinek, mintázatainak hiánya okozza, amelyeket így sem az algoritmus, sem a szakértő nem képes megtanulni.

*Többé-kevésbé szabályszerű környezetekben* (mint például az egészségügyi alkalmazások) megfelelő adatokon várhatóan jobb előrejelzéseket remélhetünk megfelelő algoritmusok alkalmazásával rövid és hosszabb előrejelzési időhorizonton egyaránt, mint szakértői becslésekre támaszkodva. Hosszabb távon a szakértők számára az jelenti a gondot, hogy nem nyílik lehetőség a tudásuk elmélyítését szolgáló, kellően nagy számú „gyakorlásra”. Az algoritmusok esetében általában az adatok elérhetősége, az adatszerkezet minősége és a modellek visszamérésének (ellenőrzésének) korlátjai jelenthetnek akadályt.

4 B bizonyos statisztikai modellek kellően robusztusak a felhasznált változók eloszlására nézve.

*Szabályszerű, jól strukturált környezetben* az algoritmikus előrejelzés elsőbbsége a kutatások alapján elvitathatatlan, feltéve, hogy a probléma szempontjából releváns, megfelelő mennyiségű és minőségű adat áll rendelkezésre. Ez esetben az algoritmikus előrejelzések lényegesen jobbak a szakértői becsléseknél. További nagy előnyük, hogy tömegszerű döntésekben automatizálva, olcsón és gyorsan alkalmazhatók. Ennek empirikus alátámasztását legjobban a tömegszerű szolgáltatást végző iparágakban (bank, biztosítás, telekom stb.) gombamód szaporodó, automatizált hitelscoreing, CRM, CHURN és egyéb statisztikai alapú döntési rendszerek esetében figyelhetjük meg. A statisztikai előrejelzések alapjául szolgáló modellek a jellemzően magasabb előrejelzési teljesítményen túl arra is alkalmasak, hogy feltárják az előrejelzés magyarázó faktorait, azok összefüggéseit egymással és az előre jelezni kívánt mennyiséggel. Ezáltal nemcsak előrejelzési céllal hasznosíthatók, hanem általuk tudásra, új ismeretekre, a modellezendő jelenség belső mechanizmusának, összefüggésrendszerének megértésére is szert tehetünk.

## IRODALOMJEGYZÉK

- ASHENFELTER, ORLEY (2007): *Predicting the quality and prices of Bordeaux wines*, AAWE working paper, No. 4.
- BAUMEISTER, ROY F. (2001): *Is there anything good about men? How cultures flourish by exploiting men*. Oxford University Press, p. 211.
- BROWN, PAUL R. (1983): Independent auditor judgment in the evaluation of internal audit functions. *Journal of Accounting Research* 21, pp. 444–455.
- CARPENTER, R. G. – GARDNER, A. – MCWEENY, P. M. – EMERY, J. L. (1977): Multistage scoring system for identifying infants at risk of unexpected death. *Arch. Dis. Childh* 53, pp. 606–612.
- CARROLL, J. S., WIENER, R. L., COATES, D., GALEGHER, J. – ALIBRIO, J. J. (1982): Evaluation, diagnosis, and prediction in parole decision-making. *Law and Society Review* 17, pp. 199–228.
- DEVAUL, R. A. – JERVEY, F. – CHAPPELL, J. A. – CARVER, P. – SHORT, B. – O'KEEFE, S. (1957): Medical school performance of initially rejected students. *Journal of the American Medical Association* 257, pp. 47–51.
- DOBELLI, ROLF (2013): *Gondolj bele, hogy ne ess bele! – A hétköznapi gondolkodás 52 csapdája*. Akadémiai Kiadó.
- GOLDING, JEAN – LIMERICK, SYLVIA – MACFARLANE, AIDAN (1985): *Sudden infant death*. Somerset: Open Books.
- HOFFMAN, PAUL J. – SLOVIC, PAUL – RORER, LEONARD G. (1968): An analysis-of-variance model for the assessment of configural cue utilization in clinical judgement. *Psychological Bulletin* 69, pp. 338–339.
- HOWARD, J. W. – DAWES, R. M. (1976): Linear prediction of marital happiness. *Personality and Social Psychology Bulletin* 2, pp. 478–480.
- KAHNEMAN, DANIEL (2013): *Gyors és lassú gondolkodás*. HVG Kiadói Rt., pp. 257–270.
- LIBBY, R. (1976): Man versus model of man: Some conflicting evidence. *Organizational Behavior and Human Performance* 16, pp. 1–12.
- LOWRY, C. (1975): *The identification of infants at high risk of early death*. Med. Stats. Report, London School of Hygiene and Tropical Medicine.
- MEEHL, P. (1954): *Clinical vs. Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*. Minneapolis, MN: University of Minnesota Press.
- SLOVIC, P. (2000): *The perception of risk*. London: Earthscan.
- SWETS, J. A. – DAWES, R. M. – MONAHAN, J. (2000): Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest* 1, pp. 1–26.
- STILLWELL, W.G. – BARRON, F. H. – EDWARDS, W. (1983): Evaluating credit applications: A validation of multiattribute utility weight elicitation techniques. *Organizational Behavior and Human Performance* 32, pp. 87–108.
- TETLOCK, PHILIP E. (2006): How accurate are your pet pundits? Project Syndicate/Institute for human sciences.
- WITTMAN, P. (1941): A Scale for Measuring Prognosis in Schizophrenic Patients. *Elgin Papers* 4, pp. 20–33.