

PREDIKTÍV MODELLEZÉS ÉS STATISZTIKAI PROFILALKOTÁS A BŰNÜLDÖZÉS SZOLGÁLATÁBAN

1. Bevezetés

A negyedik ipari forradalmat követően egyre csak nő az elérhető adatok mennyisége, megállíthatatlanul fejlődik a mesterséges intelligencia és terjed az automatizáció. Ezen tényezők kombinálásával hatékony nyomozástámogató rendszerek hozhatók létre, mint például egy automatizált profilalkotó rendszer. Egy korábbi tanulmányban¹ bemutatásra került, hogyan válhat (vagy vált, például az Egyesült Királyságban²) a kriminálpszichológia a kriminalisztika erőforrásává. A nyomok regisztrálását és szakértői gyorslemezéseket követően a begyűjtött információk alapján a hipotézisek kerülnek felállításra az elkövetővel kapcsolatban.

További adatgyűjtéseket követően pedig esetleges korrekcióra, finomításra kerül sor. Ide illeszthető be automatizált profilalkotási rendszer, ami matematikai-statisztikai módszereket használ, hogy lezárt ügyeket vizsgálva összefüggéseket keressen az elkövető és a cselekmény, illetve annak sértettje között. Ezeket az összefüggéseket használja fel a későbbiekben, amikor egy új ügy érkezik. Végezetül, amikor az lezárul, bekerül az adatbázisába és a rendszer felülvizsgálja a megtanult összefüggéseket. Előnye, hogy az érkező adatokat akár azonnal képes feldolgozni és eredményt adni, illetve a különböző hipotézisek megbízhatóságát is jelzi (például: mennyire valószínű, hogy az elkövető családi státuszára vonatkozó hipotézis helytálló).

2. Statisztikai profilalkotás, prediktív modellezés és gépi tanulás

Automatizált profilalkotási rendszer alatt gyakorlatilag a statisztikai profilalkotási eredmények és módszerek egységes, összetett alkalmazása értendő. Maga a statisztikai profilalkotás, az elkövetői profilalkotás egyik fajtája^{3,4}). A megközelítés „*elsődlegesen az elkövetői viselkedések, valamint egyéb, a bűncselekmények helyszínéből származó információ elemzését jelenti többváltozós statisztikai eljárásokkal*”⁵. Egyik célja a nyomozó tevékenységének a támogatása többek között a gyanúsított kör szűkítésével, ami jelen témához a legszorosabban kapcsolódik.

¹ Lohner Klaudia – Hermann Zsombor – Haller József: Szexuális motiváció nyomai a tetthelyen: kriminálpszichológiától a kriminalisztikáig. Belügyi Szemle 2021/10. 1813-1828 o.

² Terri Cole – Jennifer Brown: Behavioural Investigative Advice: Assistance to Investigative Decision-making in Difficult-to-detect Murder. Journal of Investigative Psychology and Offender Profiling, 2014/3. 191-220 o.

³ Lohner Klaudia: Profilalkotás – kutatásokon át a helyszíni szemléig. Belügyi Szemle, 2021/12. 2119-2135 o.

⁴ Haller József – Petőfi Attila – Mészáros Kristóf: A profilozás könnyűségéről és nehézségéről – egy perspektivikus megközelítés. Belügyi Szemle 2021/12. 2069-2086 o.

⁵ Ivaskevics Krisztián: Bűnözői profilalkotás. In Haller József (szerk.): Rendészeti pszichológia. Ludovika Egyetemi Kiadó, Budapest, 2020. 120 o.

A szóban forgó statisztikai eljárások egyik népszerű fajtája az ún. prediktív modellezés. Tág értelemben akkor alkalmazható, amikor ismert tényezők alapján ismert összefüggések segítségével szeretnénk hipotéziseket felállítani ismeretlen tényezőkről – időbeliségtől függetlenül. Jó példa az orvostudomány: diagnózis során mekkora valószínűséggel rendelkezik valaki adott betegséggel, vagy prevenció esetében mekkora valószínűséggel alakulhat ki betegség később az élete során⁶. A statisztikai profilalkotás kontextusában az ismert (bemeneti) tényezők a helyszíni nyomok, sértett jellemzői, szakértői elemzések és minden egyéb elérhető információ, ismeretlenek (kimeneti) pedig legtöbbször az elkövető jellemzői. Az összefüggések keresése során legtöbbször alkalmazott módszerek a különböző gépi tanulási algoritmusok.

A gépi tanulás (ML – *machine learning*) definíció szerint „*az a tudományterület, mely képessé teszi a számítógépeket a tanulásra, anélkül, hogy explicit programoznánk őket*”⁷. Nem szükséges tehát az összefüggések ismerete és ezeknek átadása valamelyik programozási nyelven, hanem a számítógép önállóan képes lesz ezek felismerésére és alkalmazására. Számos tudományterületre támaszkodik, mint például a statisztikára, információelméletre, valószínűségszámításra, funcionálanalízisre⁸. Nem egyenlő a mesterséges intelligenciával (MI), annak részterülete. Míg a mesterséges intelligencia egy intelligens ágens, ami kapcsolatban van a környezetével, feldolgozza a beérkező információkat és döntéseket hoz, cselekedeteket hajt végre egy cél elérése érdekében, a gépi tanulás az emberi tanulást igyekszik a lehető legjobban reprodukálni az MI szolgálatában. Az alkalmazott módszerek több csoportra oszthatók: felügyelt, felügyelet nélküli, részben felügyelt, és megerősítéses tanulás⁹. Ezek részletes és technikai bemutatása meghaladja jelen cikk kereteit, alkalmazásuk bevezetéséhez ajánlott két könyv^{10,11} az ismert programnyelv (R vagy Python) függvényében.

A következőkben bemutatásra kerülnek a prediktív modellezés céljából alkalmazott gépi tanulási módszerek a statisztikai profilalkotás területén. Nem cél az összes elért tudományos eredmény vizsgálata, hanem a felhasználhatóság szemléltetése egy-egy érdekes példán keresztül, amik a (statisztikai) profilalkotás különböző problémafelvetéseit is jól illusztrálják. A módszerek hasonlóságai alapján négy csoportra oszthatók: egyszerű predikció, összetett predikció, jellemzők csoportosítása, illetve egyének csoportosítása.

3. Egyszerű predikció

Egyszerű predikció alatt azokat a módszereket értem, ahol egy adott jellemző becslésére dolgozunk ki egy modellt. Ilyen például a klasszikus, ma már több, mint 100 éves¹² regresszió. Több ismeretlen jellemző esetében egymástól független modellek felállítása

⁶ Ewout W. Steyerberg: *Clinical Prediction Models A Practical Approach to Development, Validation, and Updating*. Statistics for Biology and Health, Springer, 2019.

⁷ Arthur L. Samuel: *Machine Learning*. The Technology Review, 1959/62. 1. 42-45 o.

⁸ Andreas Munoz: *Machine learning and optimization*. Courant Institute of Mathematical Sciences 2014/1-2, 1-14.o.

⁹ Pariwat Ongsulee: *Artificial intelligence, machine learning and deep learning*. 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), 2017, 1-6 o.

¹⁰ Gareth James – Daniela Witten – Trevor Hastie – Robert Tibshirani: *An Introduction to Statistical Learning with Applications in R*. Második kiadás. Springer, 2021.

¹¹ Gareth James – Daniela Witten – Trevor Hastie – Robert Tibshirani – Jonathan Taylor: *An Introduction to Statistical Learning with Applications in Python*. Springer, 2023.

¹² Karl Pearson: *The Life, Letters and Labors of Francis Galton*. Cambridge University Press, 1930.

lehetséges. Ezek közül általában egyik pontosabb, másik pontatlanabb eredményt hoz, illetve előfordulhat, hogy eltérő modelleknél más prediktor változók fognak kiemelkedni az adott jellemző meghatározása során. Szintén az egyszerű predikció kategóriájába esnek a fa-alapú modellek, például a RandomForest¹³ és az XGBoost¹⁴. Működésükben az emberi gondolkodásra, döntéshozásra hasonlítanak. Ha-akkor alapú döntések sorozatait állítják fel és ezek alapján próbálják meghatározni a lehetséges kimenetet.

Alkalmazásukra többször visszatérő példa az elkövető priuszának kérdése: a bűncselekmény során tanúsított viselkedésből következtethetünk-e arra, hogy korábban elkövetett-e már bűncselekményt, és ha igen, mit. Magyar mintán emberölést elkövetők esetében Ivaskevics és Almond¹⁵ kiemelte a halál okának és a sértett különböző sérüléseinek fontosságát az elkövető lehetséges priuszának megismerésében. Három tanulmány vizsgálja ugyanezt szexuális erőszak esetén^{16,17,18}. Az Egyesült Királyságban, közel 20 év különbséggel végzett vizsgálatok jelentősen eltérő eredményt mutattak, melyekre több lehetséges magyarázat felmerül, például jogi szabályozásban történt változások, módszertani eltérések, mintabeli különbségek¹⁷. Janosch és munkatársai¹⁸ az Egyesült Királyság és Spanyolország közötti különbségek bemutatásán keresztül hangsúlyozták ország-specifikus modellek létrehozásának fontosságát.

4. Összetett predikció

Összetett predikció kategóriájába sorolhatók azok a módszerek, amik egy modellben képesek kezelni az összes ismert és ismeretlen információt – a *bemeneti* és *kimeneti* változókat. A statisztikai profilalkotás területén még kevésbé elterjed Bayes-féle valószínűségi hálók¹⁹ egy jó példa erre. Röviden összefoglalva a Bayes-háló ebben a kontextusban a sértett és helyszíni jellemzők, illetve az elkövető jellemzői közötti valószínűségi kapcsolatok grafikai reprezentációja. Ezeket a kapcsolatokat felhasználva lehet kiszámolni, hogy az elkövető milyen valószínűséggel rendelkezik különböző jellemzőkkel²⁰. A Bayes-hálók számos további területen nyújthatnak segítséget a nyomozásban, ezek mutatja be Orbán²¹.

Kimondottan statisztikai profilalkotás területén emberölésekre fókuszálva Baumgartner^{22,23,24}, Stahlscmidt és munkatársai²⁵, illetve Pas²⁶ modellezte a fentebb leírt

¹³ Leo Breiman: Random Forests. Machine Learning, 2001/45, 5-32 o.

¹⁴ Tianqi Chen – Carlos Guestrin: XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, 785-794 o.

¹⁵ Krisztián Ivaskevics – Louise Almond: Predicting the criminal record of Hungarian homicide offenders from crime scene behaviours. Psychology, Crime & Law, 2019/26. 1. 22-33 o.

¹⁶ Annie Davies – Karin Wittebrood – J.L. Jackson: Predicting the criminal antecedents of a stranger rapist from his offence behaviour. Science & Justice, 1997/37, 3, 161-170 o.

¹⁷ Louise Almond – Michelle McManus – Amarat Bal – Freya O'Brien – Lee Rainbow – Mark Webb: Assisting the Investigation of Stranger Rapes: Predicting the Criminal Record of U.K. Stranger Rapists From Their Crime Scene Behaviors. Journal of Interpersonal Violence, 2018/36, 3-4, 2005-2028 o.

¹⁸ Heriberto Janosch – Carmen Jordá – Diana Nut – Susan Giles – Louise Almond: Predicting the Criminal Record of Spanish Stranger Rapists from Their Crime Scene Behaviours. Journal of Aggression, Maltreatment & Trauma, 2023/32, 1-20 o.

¹⁹ Judea Pearl: Causality: Models, Reasoning and Inference. Cambridge University Press, New York, 2000.

²⁰ Niamh Nic Daéid: Differences in offender profiling in the United States of America and the United Kingdom. Forensic Science International, 1997/90, 1-2, 25-31 o.

²¹ Orbán József: Kriminálisztikai valószínűségi becslés Bayes-hálókkal. Magyar Rendészet 2014/4. 115-130 o.

²² Kelli A. C. Baumgartner: Bayesian Network Modeling of Offender Behavior for Criminal Profiling. Master Thesis, Duke University, 2005.

összefüggéseket. Baumgartnerék arra a következtetésre jutottak, hogy ígéretes a módszer, összességében 79%-os pontosságot értek el²². Ez természetesen elkövetői jellemzőként változott: egyes esetekben pontosabb (helyszínismeret: 93,6%, nem: 87,2%), míg másoknál pontatlanabb (munkanélküliség: 46,8%, priusz-lopás: 57,4%) volt a modell. Későbbi vizsgálata során más bemeneti és kimeneti változókat figyelembe véve kollégáival²³ jelentősen jobb eredményt kaptak (90.5%). Érdekességképp összevetették szakemberek által készített profilokkal, amik esetében 62%-os átlagot mértek. Hasonlóan magas pontosságot ért el Stahlscmidt-ék²⁵ szexuális motivációjú emberölések területén. Pas²⁶ többféle tanulóalgoritmus hasonlított össze, melyek közül a legjobban teljesítő holland mintán 80% körüli pontosságot mutatott.

5. Jellemzők csoportosítása

Ebbe a csoportba azok a módszerek sorolhatók, melyek célja a különböző viselkedéses jegyek csoportosítása. Amennyiben eltérő jegyek jellemzően együtt jelennek meg, azok egy csoportba (típusba) tartoznak. Ezáltal létrehozhatók tipológiák és az új ügyben szereplő ismeretlen elkövetőről a típusba sorolást követően következtetések vonhatók le. A kapcsolódó módszerek közé tartoznak többdimenziós skálázási eljárások: SSA, főkomponens-elemzés és faktorelemzés.

David Canter és munkatársai ezen felsorolt módszerek segítségével dolgozták ki a bűncselekmények elkövetésének ún. radex-modelljét^{27,28,29,30}. Eszerint az elkövetők viselkedéses jegyei egy körre illeszthetők: középen található azok, amik szinte minden esetben megjelennek (általános viselkedések), a kör széle felé pedig a ritkán előfordulók (specifikus viselkedések) kerülnek. A közöttük levő távolságok azt reprezentálják, mennyire gyakran jelennek meg együtt. A kialakított modell alapján meghatározásra kerülnek az elkövetési tematikák, azonban ezek nem kezelhetők egymástól függetlenül. Több bűncselekmény elkövetőinél dolgoztak ki radex-modellt, melyek közé tartozik (a teljesség igénye nélkül) a szexuális erőszak³¹, gyújtogatás³², illetve sorozat-emberölés³⁴.

6. Egyének csoportosítása

Az utolsó csoportba a klaszterelemzés módszerei sorolható, melyek arra fókuszálnak, hogy a mintában található eseteket hasonlóság alapján csoportosítsák. A cél az előzőhöz

²³ Kelli A. C. Baumgartner – Silvia Ferrari – C. Gabriel Salfati: Bayesian network modeling of offender behavior for criminal profiling. Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference, 2005. 2702-2709 o.

²⁴ Kelli A. C. Baumgartner – Silvia Ferrari – George Palermo: Constructing Bayesian networks for criminal profiling from limited data. Knowledge-Based Systems, 2008/21. 7. 563-572 o.

²⁵ Stephan Stahlscmidt – Helmut Tausendteufel – Wolfgang K. Härdle: Bayesian networks for sex-related homicides: structure learning and prediction. Journal of Applied Statistics, 2013/40. 6. 1155-1171 o.

²⁶ Lotte Pas: Statistical Criminal Profiling. Master Thesis, Universiteit Leiden, 2018.

²⁷ David Canter – Rupert Heritage: A multivariate model of sexual offence behaviour: developments in offender profiling. The Journal of Forensic Psychiatry, 1990/1. 2. 185-212 o.

²⁸ David Canter – Katarina Fritzon: Differentiating arsonists: A model of firesetting actions and characteristics. Legal and Criminological Psychology, 1998/3. 1. 73-96 o.

²⁹ David Canter: Offender profiling and criminal differentiation. Legal and Criminological Psychology, 2000/5. 1. 23-46 o.

³⁰ David Canter – Laurence Alison – Emily Alison – Natalia Wentink: The Organized/Disorganized Typology of Serial Murder: Myth or Model? Psychology, Public Policy, and Law, 2004/10. 3. 293-320 o.

hasonlóan tipológia létrehozása, azonban különböző jellemzők helyett az egyéneken van a hangsúly. Két elkövető hasonlít egymásra, ha hasonló jellemzőkkel rendelkeznek, hasonló sértettek sérelmére hasonlóképp követnek el bűncselekményt. Ezekből a csoportokból jön létre a tipológia, amik összehasonlítása során megállapításra kerülnek a legfontosabb eltérések. Számos formája van, például hierarchikus, k-központú, dbscan és sok egyéb³¹.

Taylor és munkatársai³² az FBI sorozatgyilkosokra vonatkozó rendezett-rendezetlen³³, illetve Holmes és De Burger³⁴ négyes tipológiáját próbálták reprodukálni részleges sikerrel. Az FBI dichotóm tipológiájához és Canterék³⁵ eredményeihez kapcsolódva megerősítik azt az elgondolást, hogy a legtöbb sorozatgyilkos esetében a rendezetlenségre utaló jegyek mentén lehetne különbséget tenni. Emellett hangsúlyozzák, hogy „vegytiszta” tipológiák helyett vegyes tipológiákra érdemes fókuszálni.

7. Gyakorlati akadályok

A fentebb bemutatott módszerek rendelkezésre állnak és folyamatosan fejlődnek, de az automatizált statisztikai profilalkotási rendszer gyakorlati megvalósulása előtt számos leküzdendő akadály áll. Ezek közül az egyik legjelentősebb az adatok minőségének és mennyiségének a kérdése. A pontosság nagyban függ mindkettő tényezőtől. A figyelembe vett változók számának növekedésével jelentősen nő a szükséges adatmennyiség. Hagyományosan klasszifikációs problémák esetében (ahol a kimeneti változó kategorikus – legnagyobbbrészt ide tartozik a statisztikai profilalkotás) a változók számának minimum 10-szerese kell lennie az érvényes esetek számának. Hagyományos regressziónál, ez a szorzó már 50³⁶. Az idézett tanulmányokban általában 15-30 jellemzőt vizsgálnak, amihez legalább 150-300 eset, de optimális esetben, a pontosság növeléséhez ennek akár többszöröse szükséges. Minden egyes eset egy lezárt ügyet jelent adott bűncselekmény kategóriánál. Kiemelendő továbbá, hogy ezeknek mind érvényesnek kell lennie, és itt elérkezünk a második nagy akadályhoz, ami a hiányzó adatok kérdésköre.

Több kutató szembesül a hiányzó adatok jelentős számával^{37,38}. Amennyiben egy adott esetről hiányzik egy változóra vonatkozó információ, úgy az összefüggések vizsgálata során azt nem lehet figyelembe venni. Erre több megoldás létezik. Pas³⁷ a hiányzó adatokat egy új, „nem ismert” kategóriaként értelmezte, így előfordulhatott, hogy az elkövető nemére a modell azt a predikciót eredményezte, hogy „nem ismert”. Ennek a hátránya magától érthető. Klasszikus megoldás a hiányzó adatokat tartalmazó esetek kihagyása az

³¹ Mayra Z. Rodriguez – Cesar H. Comin – Dalcimar Casanova – Odemir M. Bruno – Diego R. Amacio – Luciano da F. Costa – Francisco A. Rodrigues: Clustering algorithms: A comparative approach. PLoS ONE, 2019/14. 1. 1-34. o.

³² Sandie Taylor – Daniel Lambeth – Georga Green – Rachel Bone – Marie A. Cahillane: Cluster Analysis Examination of Serial Killer Profiling Categories: A Bottom-Up Approach. Journal of Investigative Psychology and Offender Profiling, 2012/9. 1. 30-51. o.

³³ Robert K. Ressler – Ann W. Burgess – John E. Douglas: Sexual Homicide: Patterns and Motives. Free Press, 1998.

³⁴ Ronald M. Holmes – James E. DeBurger: Profiles in Terror - The Serial Murderer. Federal Probation, 1985/49. 1. 29-34 o.

³⁵ Lásd: 30. lj.

³⁶ Valliappa Lakshmanan – Sara Robinson – Michael Munn: Machine Learning Design Patterns - Solutions to Common Challenges in Data Preparation, Model Building, and MLOps. O'Reilly Media, 2020.

³⁷ Lásd: 28. lj.

³⁸ Lásd: 1. lj.

elemzésből, azonban ez is torzítani tud az eredményeken³⁹. Egy harmadik lehetőség az ún. imputáció, azaz adatpótlás. Ennek során megpróbáljuk a hiányzó adatot mesterségesen helyettesíteni. Több fokozata létezik, melyek közül a leginkább összetett és ígéretes a többszörös imputáció⁴⁰. A módszerek adottak, azonban nem született olyan kutatás, ami a profilalkotás területén ezek összehasonlítását pontosságuk alapján.

Az adatok mennyisége mellett azok minősége mellett sem lehet elmenni. Az adatok minősége alatt azt értem, hogy egyrészt mindegyik megfelel a valóságnak, másrészt adott változó esetében minden kitöltő ugyanazt értette a kérdés alatt. Az elkövetői profilalkotás során alkalmazott adatbázisok jellemzően egy adatlapra épülnek, amit nyomozók/szakemberek töltenek ki, és több ellenőrzést követően kerülnek csak be az adatok az adatbázisba. Ez így van hazánkban is⁴¹, és a nemzetközileg elismert VICLAS⁴² esetében is. A VICLAS adatlap megbízhatóságát két kutatásban vizsgálták, melyek eredménye kérdéseket vet fel. Azt találták, hogy a kérdőívet kitöltők közötti egyetértés (százalékban mérten) eltérő (79,3%-87,7% bűncselekménytípusonként⁴³, illetve 36%-62,87% kérdésenként⁴⁴).

További akadály a vizsgált változók kérdésköre. Az adatok mellett ez a második legjelentősebb szempont, amit a statisztikai (és bármilyen) profilalkotás során figyelembe kell venni. A „felesleges” (adott kontextusban nem releváns) jellemzőket azonosítani és kizárni statisztikai módszerekkel lehetséges, azonban egy fontos változó hiánya pótolhatatlan és torzít az eredményen.

Végezetül pedig meg kell említeni a fekete-doboz jelenséget. Gépi tanulási módszereinek alkalmazása során jellemzően kettős a fókusz: a kapott eredmény mennyire pontos, illetve a vizsgált jellemzők között milyen összefüggéseket lehet találni. Sok esetben szembesülni kell azzal, hogy az egyik javára felborul az egyensúly. A neurális hálónál az eredmény nagyon pontos tud lenni, azonban az összefüggések közvetlenül nem hozzáférhetőek az emberek számára⁴⁵. Az algoritmus tanulta meg ezeket helyettünk. Hosszú és bonyolult elemzésekre van szükség, hogy feltárható legyen, a végeredmény kiszámításának pontos módja.

8. Limitációk

A statisztikai profilalkotás során elengedhetetlen a szaktudás, szakember, ami/aki kiindulópontként szolgál és értelmezi az eredményeket. A gépi tanulási módszerek eredményei hasznos információkkal szolgálnak, azonban önmagukban nem elegendők. Szükséges az adott ügy mélyreható ismerete, a többi módszer eredményeinek

³⁹ John W. Graham: Missing data analysis: making it work in the real world. *Annual Review of Psychology*, 2009/60. 549-576 o.

⁴⁰ Stef van Buuren: *Flexible Imputation of Missing Data*. Chapman & Hall, 2018.

⁴¹ Lásd: 1. l.j.

⁴² Peter I. Collins – Gregory F. Johnson – Alberto Choy – Keith T. Davidson – Ronald E. Mackay: Advances in violent crime analysis and law enforcement: The Canadian violent crime linkage analysis system. *Journal of Government Information*, 1998/25. 3. 277-284 o.

⁴³ Melissa Martineau – Shevaun Corey: Investigating the Reliability of the Violent Crime Linkage Analysis System (ViCLAS) Crime Report. *Journal of Police and Criminal Psychology*, 2008/23. 51-60 o.

⁴⁴ Brent Snook – Kirk Luther – John C. House – Craig Bennell – Paul J. Taylor: The Violent Crime Linkage Analysis System: A test of Interrater Reliability. *Criminal Justice and Behavior*, 2012/39. 5. 607-619 o.

⁴⁵ David Castelvecchi: The black box of AI. *Nature*, 2016/538. 20-23 o.

összehasonlítása, hogy végezetül megalkotható legyen az elkövető profilja. A profilalkotás sikerének kulcsa a multidiszciplinaritás, melyről Lohner⁴⁶ ír részletesen.

9. Összegzés

A gépi tanuláson alapuló automatizált profilalkotási rendszer elérhető távolságban van. A nemzetközi és hazai szakirodalomban elért eredmények, pontosságok azt mutatják, hogy érdemes tovább kutatni és fejleszteni a módszereket. Az előrelépés előtt álló legnagyobb akadály az adatok mennyiségének és minőségének kérdésköre. A problémafelvetés és a módszerek adottak, minden a megfelelő adatbázis létrehozásán áll.

A nemzetközi eredmények alapján a statisztikai profilalkotással kapcsolatban több következtetés vonható le. Egyrészt kultúránként eltérőek az összefüggések, így ország-specifikus modellek létrehozása adhat pontos eredményt. Másrészt felvetül a lehetősége annak, hogy az összefüggések az idő múlásával is változnak, ez a kérdéskör további kutatást igényel. Harmadrészt számos módszer érhető el, melyek egyes esetekben, egyes jellemzőknél pontosabbak, míg másoknál kevésbé. Szükséges több modell felállítása, majd ezek összehasonlítása. Végezetül pedig a statisztikai profilalkotás önmagában nem elegendő. A feltett kérdéshez nyomozóra, pszichológusra, szakértőre van szükség, és az eredményeket sem lehet nélkülük értelmezni. A statisztikai profilalkotás egy olyan eszköz, ami segíthet a prioritizálásban, információval szolgálhat döntéshozatalhoz, azonban ezeket helyettesíteni nem tudja.

⁴⁶ Lásd: 3.