



ÓBUDAI EGYETEM
ÓBUDA UNIVERSITY



Mérnöki Szimpózium a Bánkiban
(ESB 2024)



Mesterséges intelligencia alapú döntéstámogató rendszer fejlesztése gépjármű biztosítási csalások feltárásához – 3. rész

AI-based decision support system for vehicle insurance fraud detection – Part III.

¹Matyasovszki Márton Gábor, ²Váradi Péter, ³Lukács Judit

¹ Óbudai Egyetem, Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar, Budapest, Magyarország, wmpc0r@stud.uni-obuda.hu

² Óbudai Egyetem, Biztonságtudományi Doktori Iskola, Budapest, Magyarország, varadi.peter@phd.uni-obuda.hu

³ Óbudai Egyetem, Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar, Budapest, Magyarország, lukacs.judit@bkgk.uni-obuda.hu

Összefoglalás

A biztosítási csalások világszerte ismert gazdasági bűncselekménynek számítanak. Súlyos károkat okoznak, valamint a biztosító társaságok és ügyfelek közötti bizalmat is rombolják. Jelen publikáció keretén belül egy olyan döntéstámogató rendszert mutatunk be, mely alkalmas lehet a gyanús esetek kiválogatására a kötelező gépjármű felelősségbiztosítási piacon már a kivizsgálás igen korai szakaszában. Ennek megvalósítására valós kártérítési esetek adatait alkalmaztuk. Létrehoztunk egy hierarchikus fuzzy struktúrát, melyben az okozó és a károsult oldalán egyaránt meghatároztuk a csalás kockázati szintjét, majd ezeket összevonva számítottuk ki az eset teljes kockázatát. Megállapítottuk, hogy a küszöbérték felett részletesebb vizsgálat szükséges képzett kárszakértő részéről. Összességében elmondható, hogy az általunk létrehozott rendszer alkalmas lehet az esetek kiválogatására, a szakértelem fókuszálásával pedig a csalásnyűs esetek alaposabb áttekintése által hatékonyabbá tehető a kártérítési folyamat. A közlemény harmadik része a megalkotott rendszert ismerteti.

Kulcsszavak: biztosítási csalás, mesterséges intelligencia, fuzzy logika

Abstract

Insurance fraud, characterized by false or exaggerated claims, is a significant economic crime worldwide, undermining the trust between insurance companies and their customers. Detecting such fraud is crucial. This study introduces a decision support system for the early identification of suspicious claims in the compulsory motor liability insurance market. We used real claim data involving two privately owned passenger cars in cases where no personal injury occurred. Based on this, we developed a hierarchical fuzzy structure to assess the fraud risk for both parties involved (the policyholder and the claimant), resulting in an overall case risk. The results indicate that only cases exceeding a certain threshold need to be further examined by qualified adjusters. This system enhances efficiency by enabling targeted reviews of potential fraud cases. The third part represents the created decision-support system.

Keywords: insurance fraud, artificial intelligence, fuzzy logic

1. Bevezetés

A közlemény előző részeiben ismertettük a magyarországi és nemzetközi helyzetet a gépjármű biztosítási ágazatban tapasztalt csalásyanús esetek kapcsán.

Ezt követően bemutatásra került a rendelkezésre álló adatbázis és a fuzzy logika. Megállapítottuk, hogy ez a módszertan kifejezetten alkalmas az emberi gondolkodásmód reprezentálására [1]. Ezen felül, a Mamdani-típusú következtetés [2] jól használható szakértői tudás alapú rendszerek megalkotásához.

2. Adaptálás

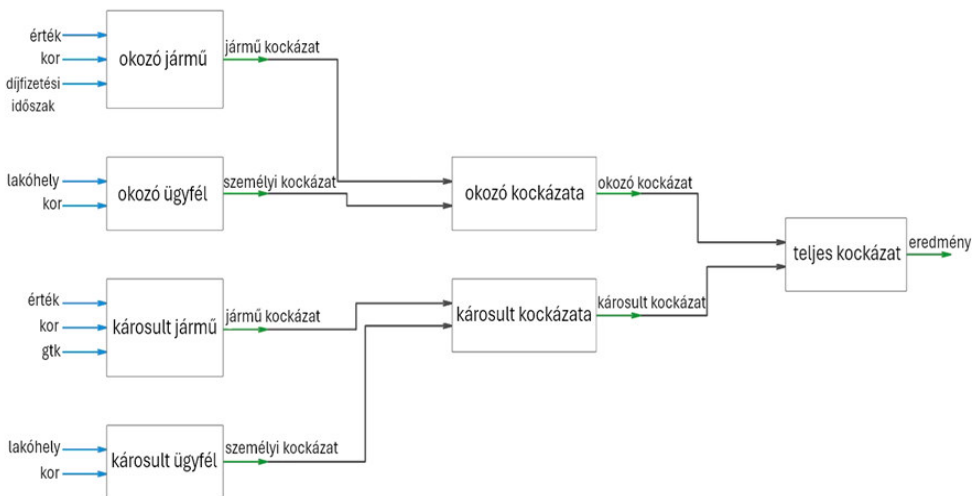
Az adaptálás során egy olyan szabály-alapú döntési modellt építettünk, amely képes a biztosítási csalások azonosítására és a különböző kockázati tényezők elemzésére. A modell kockázati besorolások alapján előrejelzést nyújt a csalásyanús kárigényekkel kapcsolatban.

A feladat megoldása során elsődlegesen az volt a cél, hogy a rendszer rugalmasságát és gyorsaságát megőrizzük. Ehhez szem előtt kellett tartani a szabálybázis strukturális kialakítását. Ugyanis számos bemeneti paramétert kellene figyelembe venni, melyek esetén a potenciális paraméterkombinációk száma jelentősen növekszik. Ez a szabálybázis méretének növekedését eredményezi, ami gyakorlatilag igen jelentős mértékű számítási kapacitás növekedést eredményezne.

Emiatt egyszerűbb következtető rendszerek hierarchikus összekapcsolása által létrehozható fuzzy fa kialakítása mellett döntöttünk, mely lehetővé teszi sok bemeneti független változó esetén is a szabálybázis méretének kordában tartását.

2.1 Fa struktúra bemutatása

A hierarchikus struktúra (lásd 1. ábra) első szintjén a független bemeneti paraméterek találhatók, a további szinteken pedig ezek részletesebb csoportosítása és a kapcsolódó kockázati szintek megállapítása és összegzése történik.



1. ábra A modell struktúrája

2.2 Bemeneti paraméterek

Az első szinten a káresetben érintett ügyfelek és járművek tulajdonságai szerepelnek. Ezeket két fő csoportra osztottuk: tárgyi (gépjárműhöz kapcsolódó) és személyi tényezők. Az 1. táblázat tartalmazza a modellben felhasznált mennyiségi változók szintjeinek jellemző értékeit és megnevezéseit. Melyeket okozói és károsult oldalon egyaránt vizsgáltunk.

1. táblázat Független változók és nyelvi leírásuk

Tárgyi paraméterek		Személyi paraméterek	
<i>Jármű életkora, év</i>	Nyelvi változó	<i>Születési év</i>	Nyelvi változó
0-2	új	-1964	Baby Boomer
3-7	fiatal	1965-1980	X generáció
8-14	közepesen idős	1981-1995	Y generáció
15+	régi	1996-	Z generáció
<i>Jármű értéke, Ft</i>	Nyelvi változó	<i>Lakosság, fő</i>	Nyelvi változó
1-500 000	nagyon alacsony	1-20 000	község-kisváros
500 001-1 600 000	alacsony	20 001-100 000	középváros
1 600 000-3 000 000	közepes	100 001-1 000 000	nagyváros
3 000 000+	drága	1 000 000+	metropolisz

Ezen felül további két, minőségi változót is beépítettünk a modellbe: az okozó oldali tárgyi kockázatnál megjelenik a biztosítási szerződés *díjfizetési időszaka*, mely lehet negyedéves, féléves vagy éves bontású, illetve károsult oldalon a *gazdasági totálkár* (GTK) bekövetkezésének ténye (igen, nem).

2.3 Szabálybázis létrehozása

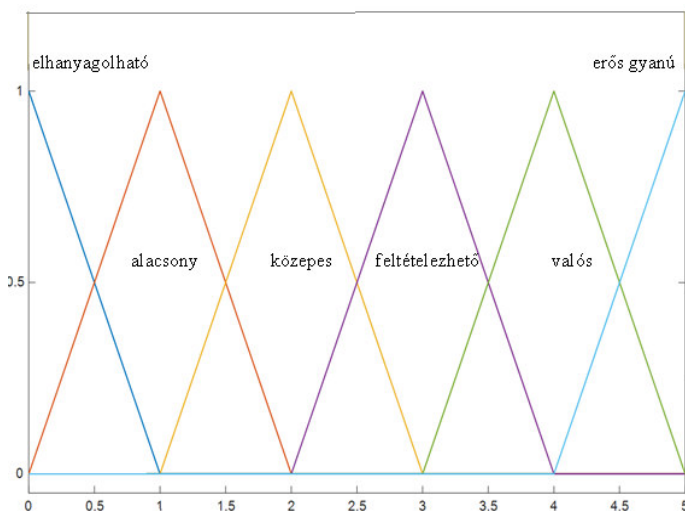
A fuzzy következtető rendszer döntési alapját a szabálybázis adja, amely a kockázatok különböző szintekre történő besorolásáért felel. A fuzzy fa első szintjének kimeneti oldalán meghatároztuk okozói és károsult oldalon a tárgyi, valamint a személyi kockázatot. Ezek partíciója

6 különböző szintet tartalmazott: elhanyagolható, alacsony, közepes, feltételezhető, valós és erős gyanú. A besorolás alapján az adott paraméterkombináció esetén az adatbázis alapján tapasztalt biztosítási csalás előfordulási aránya alapján határoztuk meg.

A fuzzy fa következő szintjén oldalanként összevonásra került a tárgyi és személyi kockázatok értéke. A szabálybázisban alkalmazott összefüggéseket a 2. táblázat alapján adtuk meg, ahol az egyes színek az 2. ábrán látható tagsági függvényekre utalnak. Itt a részkimenet egy 0...5 közé eső kockázati szint érték.

2. táblázat Okozói kockázat szabálybázisa

Okozói kockázat		Tárgyi kockázat					
		elhanyagolható	alacsony	közepes	feltételezhető	valós	erős gyanú
Személyi kockázat	elhanyagolható						
	alacsony						
	közepes						
	feltételezhető						
	valós						
	erős gyanú						



2. ábra Tagsági függvények

Végezetül, a rendszer harmadik szintjén történt az okozó és károsult oldai kockázatok összevonása, mely az adott kárigényhez tartozó teljes kockázati szintet adta, szintén 0...5 közé eső skálán. A magasabb kockázati szint erőteljesebb csalásgyanút jelent. Az alkalmazott összefüggéseket a 3. táblázat tartalmazza. A partíció elemei és szintjei szintén az 2. ábra szerint kerültek megadásra.

3. táblázat Teljes kockázat szabálybázisa

Teljes kockázat		Károsult kockázata				
		elhanyagolható	alacsony	közepes	feltételezhető	erős gyanú
Okozó kockázata	elhanyagolható					
	alacsony					
	közepes					
	feltételezhető					
	valós					
	erős gyanú					

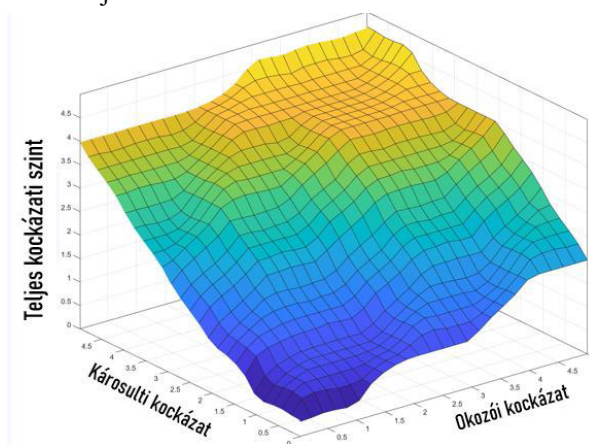
2.4 Defuzzifikációs módszer

A fuzzy logikai rendszerek esetében a defuzzifikáció egy lényeges lépés, amely során a fuzzy halmazokat éles, crisp értékekké alakítjuk [3]. Ennek célja, hogy a rendszer által számított kockázati szint egyszerűen értelmezhető legyen. Mi két defuzzifikációs módszert, a LOM (Last of Maxima) és a centroid (Center of Gravity) módszereket teszteltük.

3. Eredmények

Matematikai modell pontosságának, meggyfelelőségének vizsgálatát a létrehozásához használt adatbázison felhasználásával hajtottuk végre. Mind a 4000 eset vonatkozásában meghatároztuk a létrehozott fuzzy fa átla számított teljes kockázati szint értékét. Emellett külön elvégeztük a kalkulációkat a centroid és a LOM defuzzifikációs technikák alkalmazásával. Ezeket az értékeket összehasonlítottuk a biztosító csalásra utaló döntésével.

Arra a következtetésre jutottunk, hogy a centroid eljárás nagyobb pontosságot eredményez, így a további elemzéseket ennél a módszernél hajtottuk végre. A teljes kockázati szintet ábrázolja a 3. ábra az okozói és a károsulti kockázat függvényében, ahol az alacsonyabb értékek késsel, a magasabbak sárgával vannak jelölve.



3. ábra Teljes kockázati szint alakulása

Az összehasonlítások során a küszöbértéket 2 értékű teljes kockázatban határoztuk meg az alábbiak szerint: ezen érték alatt a csalás gyanúja kellően alacsonynak tekinthető, nincs szükség az eset részletesebb vizsgálatára. Azonban a 2...5 közé eső értékek fokozatosan ugyan, de egyre

magasabb rizikót jelentenek, itt részletes, átfogó vizsgálatokra van szükség a család gyanújának megerősítésére vagy annak eloszlatására.

Az elemzések alapján, és a 2, mint küszöbérték alkalmazása mellett, a modell 72%-os egyezést mutatott a biztosító által hozott döntésekkel. Ez azt jelenti, hogy közel 2900 esetben hatékonyan sikerült az alacsony kockázatú esetek kiválasztása. Mindez lehetőséget nyújt az erőforrások hatékony átcsoportosítására, így kiemelt fókusz mellett kezelhető a fennmaradó, magasabb kockázatú kárigények csoportja. Mindez elősegítheti a hatékonyabb családetektálást, továbbá az erőforrások átcsoportosításával az ügyintézés gyorsulhat, így csökkenthető a kárigények átfutási ideje.

4. Konklúzió

Jelen tanulmány keretein belül bemutatásra került egy három szintű fuzzy fa, melynek célja a gépjármű biztosítási családok kiszűrése volt. A modellalkotás során az elmúlt 10 évből származó valós kárigényeket használtunk. Elkülönítve vizsgáltuk az okozói és a károsulti oldalon egyaránt az érintett szerződésekhez tartozó járművek és személyek hatását. Ezek összevont elemzésére Mamdani-típusú következtető rendszereket kapcsoltunk hierarchikus struktúrába. Végezetül meghatároztuk az egyes esetekhez tartozó teljes kockázati szintet, melyet a 0...5 skálán értelmeztünk.

Az eredmények alapján az alábbi következtetésekre jutottunk:

- Megállapítottuk, hogy a bemutatott 10 elemű fuzzy fa alkalmas a kárigények kockázati szintjének meghatározására.
- Kimutattuk, hogy a legjobb teljesítmény a centroid defuzzifikációs eljárás alkalmazása esetén érhető el.
- Kijelenthető, hogy a rendszer segítséget nyújthat a kárrendezési folyamat során az emberi erőforrások hatékony csopontosítására.

Összességében tehát megállapítható, hogy a bemutatott modell egy hatékony eszköz lehet a gépjármű biztosítási családok kiszűrésére, mely pénzt és időt takarít meg a biztosítótársaságoknak és gyorsabb, pontosabb ügyintéztést biztosíthat a lakossági ügyfelek számára.

Köszönetnyilvánítás

A Kulturális és Innovációs Minisztérium 2024-2.1.1 kódszámú Egyetemi Kutatói Ösztöndíj Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

5. Hivatkozások

- [1] Zadeh, L. A. (1965). Information and control. Fuzzy sets, 8(3), 338-353.
- [2] Mamdani, E. H. (1974, december). Application of fuzzy algorithms for control of simple dynamic plant. In Proceedings of the institution of electrical engineers (Vol. 121, No. 12, pp. 1585-1588). IET
<https://doi.org/10.1049/piee.1974.0328>
- [3] Yager, R. R., Filev, D. P. (1994). Essentials of fuzzy modeling and control. New York, 388.