



Mesterséges intelligencia alapú döntéstámogató rendszer fejlesztése gépjármű biztosítási csalások feltárásához – 1. rész

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

¹Berzicai Kevin, ²Váradi Péter, ³Lukács Judit

¹ Óbudai Egyetem, Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar, Budapest, Magyarország, kevinberzicai@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@bgk.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 csalásgyanús esetek alaposabb áttekintése által hatékonyabbá tehető a kártérítési folyamat. A közlemény első része a biztosítási csalásokhoz kapcsolódó kutatások sajátosságait, valamint az esetek jellemzőit 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 first part introduces the insurance fraud-related studies and the characteristics of the revealed fraudulent cases.

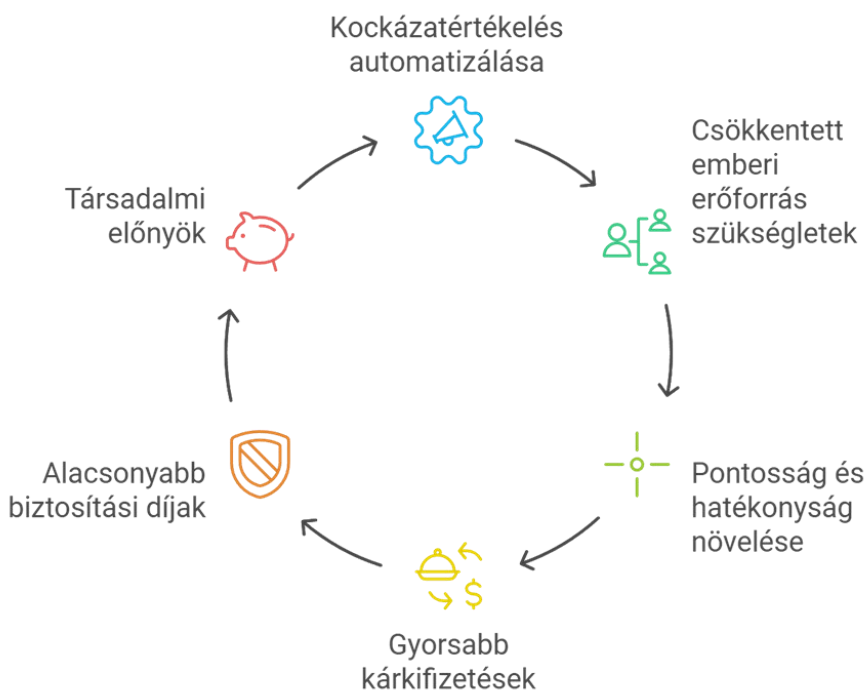
Keywords: insurance fraud, artificial intelligence, fuzzy logic

1. Bevezetés

A biztosítási csalások magas előfordulása és az általuk okozott pénzügyi károk komoly problémát jelentenek a biztosítótársaságok és ügyfeleik számára. Ez a probléma nemcsak a biztosítók nyereségességét veszélyeztetheti, hanem a társadalom egészét is, hiszen a csalások más tényezők mellett közvetlen hatással vannak a biztosítási díjak alakulására, és bizalmatlanságot kelthetnek az ügyfelek körében. A mesterséges intelligencia gyors fejlődése lehetőséget kínál a biztosítóknak arra, hogy hatékonyabb módszerekkel lépjenek fel a csalások ellen.

A KSH adatai szerint Magyarországon évente 15-18 ezer közúti baleset történik, amelyek közül feltehetően 14%-a azaz nagyjából 2100-2500 olyan biztosítási csalás történik évente, amiket a biztosító társaságok nem küszöböltek ki. Az e-kárbejelentő rendszerek megjelenésével a kárbejelentési folyamatok egyszerűsödtek és felgyorsultak, azonban a digitális csatornák új lehetőségeket is nyitnak a csalók számára [1][2].

A mesterséges intelligencia alapú detektálás automatizálhatja és javíthatja a kockázatértékelési folyamatokat (lásd 1. ábra), csökkentve az emberi erőforrás szükségességét az egyértelmű káresemények esetében, ezáltal növelve a pontosságot és a hatékonyságot a kevésbé egyértelmű csalárd eseteknél.



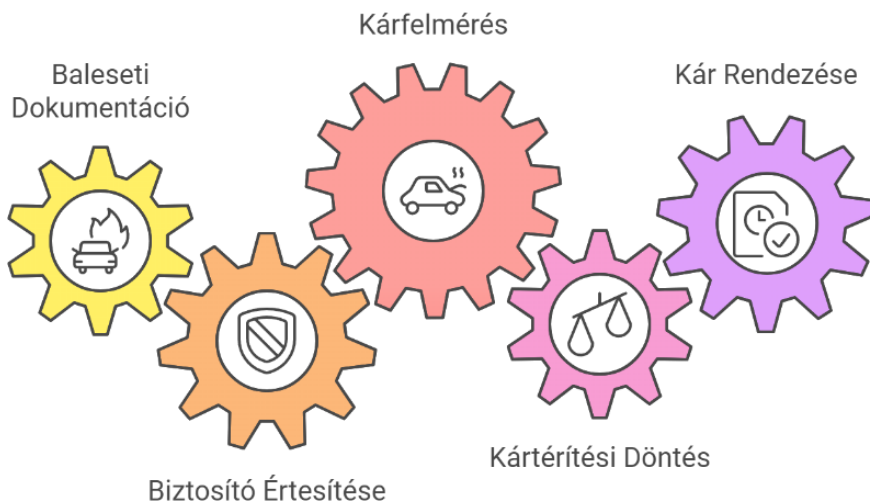
1. ábra Hatékony biztosítási kárrendezés ideális ciklusa

A folyamat felgyorsulása azt is jelenti, hogy a kifizetésre jogosultak hamarabb jutnak az őket illető összeghez úgy, hogy a biztosítási díjak kisebb mértékben emelkednek a biztosítók csökkent kiadásainak hatására.

Összegezve a téma napjainkban egyre fontosabb, mivel a biztosítási csalások elleni küzdelem nemcsak a biztosítók, hanem az ügyfelek és a társadalom közös érdeke is. A kutatás motivációja az,

hogy bemutassuk, hogyan alkalmazhatóak a mesterséges intelligencia nyújtotta technológiák ezen a területen, ezzel segítve a biztosítási rendszer gördülékeny működését.

Hazánkban a KGFB-hez (Kötelező Gépjármű Felelősség Biztosításhoz) kapcsolódó kárrendezési folyamat a 2. ábrán látható lépésekből áll.



2. ábra Kárrendezés folyamata Magyarországon

Gyorsabb ügyintézés, az online bejelentés a legtöbb biztosító esetében jelentősen felgyorsítja a folyamatot, mert az adatokat automatikusan továbbítják a rendszerbe, minimalizálva az emberi hiba lehetőségét.

A digitális csatornák használatával azonban új biztonsági kérdések is felmerülnek, például az adatvédelem és a digitális csalások elleni védelem szükségessége. Az e-kárbejelentő rendszerek adatainak helytelen kezelése vagy hamis bejelentések könnyen vezethetnek csalási esetekhez, ezért a biztosítók egyre több erőforrást fordítanak a mesterséges intelligencia alapú csalásfelderítési módszerek fejlesztésére.

A környező országok kárbejelentési rendszerei hasonlóan működnek a magyarországiakhoz, azonban országonként eltérhetnek a digitalizáció mértékében, a kárbejelentési folyamat automatizáltságában, a biztosítási rendszer berendezkedésében és a csalások kiszűrésének módszereiben.

A biztosítási csalás detektálása bináris osztályozási vagy többszörös osztályozási problémaként is kezelhető. Számos kutató alkalmazott gépi tanulási modelleket az autós biztosítási csalások detektálására, és hasznos eredményeket értek el. Viaene és társai [4], Kaščelan és társai [5], valamint Li és társai [6] a Bayes-i modellezés, klaszterelemzés, adatbányászat és véletlen erdő teljesítményét vizsgálták a gépjármű biztosítási csalás detektálásában. David és társai [7] a biztosított életkor változásának elemzésével a csalárd viselkedés magas kockázatú populációjának jellemzőit és tulajdonságait tárta fel. He és társai [8], Guo és társai [9] és Wang és társai [10] tovább kutatták a mélytanulási modellek potenciálját a csalás detektálásában. Subudhi és társai [11] és Majhi és társai [12] hibrid modelleket építettek, amelyek hatékonyan képesek detektálni a gépjármű biztosítási csalásokat. Tuo és társai [13] és Liu és társai [14] a biztosítási csalásokat játékelméleti megközelítésből vizsgálták Kínában. Zhao és társai [15], Tang és társai [16] és Wang és társai [17]

szintén gépi tanulási módszereket alkalmaztak a biztosítási csalás gyanús viselkedés modellezésére a kínai gépjármű biztosítási piac adatai alapján. Csak nemrégiben kezdtek el Yan és társai [18][19], Yu és társai [20] és Yang és társai [21] mélytanulási modellek és kevert modellek alkalmazásával biztosítási csalási problémákat elemezni, és előrehaladást értek el az autós biztosítási csalás detektálása terén.

Kiemelendő továbbá, hogy a csalásfelderítő rendszerek esetén nem elég a teljesítmény vizsgálta, hanem szükséges a költséghatékonyság elemzése is a hosszútávú fenntarthatóság érdekében. Ez a megközelítés számos egyéb tudományterületen is kiemelt kutatási terület, különösen az egészségügyben [22][23]. Ezen felül Benedek és Nagy [24] egy költségalapú összehasonlítást mutat be, arra a következtetésre jutva, hogy a vizsgált hagyományos és Mesterséges intelligencia alapú rendszerek között nem mutatható ki egyértelműen az MI jobb teljesítménye.

Sajnos azonban Magyarországon nincs hatékony, mesterséges intelligencián alapuló „jó módszer” a biztosítási csalások kiszűrésére [24]. Kijelenthető, hogy a fuzzy alapú döntéstámogató rendszerek a csalásdetektálás területén biztató eredményeket nyújthatnak. Ennek a módszertannak a vizsgálatát mutatja be Váradi és társai [25][26], ahol már egy elméleti döntéstámogató rendszer megalkotása is megjelenik [27]. Mindezek alapján megállapítható, hogy a gépjármű biztosítási csalások feltárása és kiszűrése fontos feladat mind hazai, mind nemzetközi szinten.

Jelen tanulmány keretein belül egy hierarchikus felépítésű, többlépcsős fuzzy fa kerül bemutatásra, mely alkalmas a kárigények csalás gyanús jellegének kockázatainak előrejelzésére. A rendszer valós esetek alapján került megalkotásra. Egyszerűen meghatározható változók segítségével vizsgálja a balesetben érintett személyek és járművek kapcsán tapasztalható kockázat meghatározására, melyet okozói és károsulti oldalunk egyaránt vizsgálunk. Végezetül a kárigénylési folyamat teljes kockázati szintje kerül meghatározásra, mely utal az eset rizikójára is. A létrehozott fuzzy fa kifejezetten alkalmas a biztosító oldaláról szükséges emberi erőforrások és szakértelem optimális elosztására, így elősegítheti a kárrendezési eljárás felgyorsulását.

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.

2. Hivatkozások

- [1] https://www.ksh.hu/docs/hun/xstadat/xstadat_evkozi/e_feb002.html (Megtekintve: 2024. október 12.)
- [2] <https://www.autoklub.hu/klub/hirek/baleseti-statisztikak/> (Megtekintve: 2024. október 12.)
- [3] <https://www.insurancefraud.org> (Megtekintve: 2024. október 14.)
- [4] Viaene, S., Dedene, G., Derrig, R.A. (2005). Auto claim fraud detection using Bayesian learning neural networks. *Expert Systems with Applications* 29(3), 653-666 <https://doi.org/10.1016/j.eswa.2005.04.030>
- [5] Kaščelan, L., Kaščelan, V., & Novović-Burić, M. (2016). A data mining approach for risk assessment in car insurance: evidence from Montenegro. In *Big Data: Concepts, Methodologies, Tools, and Applications* (pp. 2028-2046). IGI Global. <https://doi.org/10.4018/978-1-4666-9840-6.ch094>

- [6] Li, Y., Yan, C., Liu, W., & Li, M. (2018). A principle component analysis-based random forest with the potential nearest neighbor method for automobile insurance fraud identification. *Applied Soft Computing*, 70, 1000-1009. <https://doi.org/10.1016/j.asoc.2017.07.027>
- [7] David, M., & Jemna, D. V. (2015). Modeling the frequency of auto insurance claims by means of poisson and negative binomial models. *Analele stiintifice ale Universitatii "Al. I. Cuza" din Iasi. Stiinte economice/Scientific Annals of the "Alexandru Ioan Cuza" University of Iasi Economic Sciences* 62 (2), 2015, 151-168 <https://doi.org/10.1515/aicue-2015-0011>
- [8] He, X., & Chua, T. S. (2017, August). Neural factorization machines for sparse predictive analytics. In *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 355-364). <https://doi.org/10.1145/3077136.3080777>
- [9] Guo, J., Liu, G., Zuo, Y., Wu, J. (2018, November) Learning sequential behavior representations for fraud detection. In *2018 IEEE International Conference on Data Mining (ICDM)*, pp. 127-136 <https://doi.org/10.1109/ICDM.2018.00028>
- [10] Wang, Y., & Xu, W. (2018). Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decision Support Systems*, 105, 87-95. <https://doi.org/10.1016/j.dss.2017.11.001>
- [11] Subudhi, S., & Panigrahi, S. (2020). Use of optimized Fuzzy C-Means clustering and supervised classifiers for automobile insurance fraud detection. *Journal of King Saud University-Computer and Information Sciences*, 32(5), 568-575. <https://doi.org/10.1016/j.jksuci.2017.09.010>
- [12] Majhi, S. K. (2021). Fuzzy clustering algorithm based on modified whale optimization algorithm for automobile insurance fraud detection. *Evolutionary Intelligence*, 14(1), 35-46. <https://doi.org/10.1007/s12065-019-00260-3>
- [13] Tuo, G., Duan, J. (1999). Game Theory Analysis of Insurance Fraud. *Journal of Capital University of Economics and Business* (3), 51-54.
- [14] Liu, X., & Jin, J. (2004). The Insurance Fraud Game and Insurance Contract Based on Optimal Game Strategies. *Systems Engineering—Theory & Practice*, 24(2), 19-24.
- [15] Zhao, G., & Wu, H. (2010). Is There Moral Hazard in Chinese Automobile Insurance Market?—Evidence from Dynamic Renewal Policies. *Journal of Financial Research*, 6, 175-188.
- [16] Tang, J., & Mo, Y. (2013). Construction of auto insurance anti-fraud system based on data mining technology. *Shanghai insurance*, 11, 39-42.
- [17] Wang, H. (2016). A Research on Chinese Insurers' Moral Hazard Screening in Operation: From the Big Data Hadoop Clustering Analysis Technology Perspective. *Insurance Studies*, 2, 59-67.
- [18] Yan, C., Li, Y., & Sun, H. (2017). A Research on Automobile Insurance Fraud Identification Based on Random Forest Model and Ant Colony Optimization Algorithm. *Insurance Studies*, 6, 114-127.

- [19] Yan, C., Li, M., & Zhou, X. (2019). Improved genetic algorithm for vehicle insurance fraud identification model based on BP neural network. *Journal of Shandong University of Science and Technology (Natural Science)*, 38(5), 72-80.
- [20] Yu, W., Feng, G. F., & Zhang, W. J. (2017). Research on fraud detection system and gang identification of motor vehicle insurance. *Insurance Research*, 2017(2), 63-73.
- [21] Yang, J., Chen, K., Ding, K., Na, C., & Wang, M. (2023). Auto insurance fraud detection with multimodal learning. *Data Intelligence*, 5(2), 388-412. https://doi.org/10.1162/dint_a_00191
- [22] Lee, A., Taylor, P., Kalpathy-Cramer, J., & Tufail, A. (2017). Machine learning has arrived!. *Ophthalmology*, 124(12), 1726-1728. <https://doi.org/10.1016/j.ophtha.2017.08.046>
- [23] Hill, N. R., Sandler, B., Mokgokong, R., Lister, S., Ward, T., Boyce, R., ... & Gordon, J. (2020). Cost-effectiveness of targeted screening for the identification of patients with atrial fibrillation: evaluation of a machine learning risk prediction algorithm. *Journal of medical economics*, 23(4), 386-393. <https://doi.org/10.1080/13696998.2019.1706543>
- [24] Benedek, B., Nagy, B. Hagymányos versus MI-alapú csalásfelderítés: Költséghatékonyság a gépjármű-biztosítások területén. *HITELINTÉZETI SZEMLE*, 77. <https://doi.org/10.25201/HSZ.22.2.77>
- [25] Váradi, P., Lukács, J., & Horváth, R. (2023, May). Examination of vehicle fraud detection possibilities with the help of Fuzzy inference system. In *2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI)* (pp. 000353-000358). IEEE. <https://doi.org/10.1109/SACI58269.2023.10158631>
- [26] Váradi, P., Lukács, J., & Horváth, R. (2024, April). Response Surface Methodology to Characterize the Indicators for the Early Detection of Fraudulent Cases in the Motor Insurance Market. In *2024 IEEE 11th International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC)* (pp. 000201-000204). IEEE. <https://doi.org/10.1109/ICCC62278.2024.10582971>
- [27] Váradi, P., Lukács, J., & Horváth, R. (2024). Elméleti fuzzy következtetési rendszer a csalásgyanús esetek korai felismerésére a gépjármű-biztosítási ágazatban. *Közlekedéstudományi Szemle*, 74(5), 31-43. <https://doi.org/10.24228/KTSZ.2024.5.3>