

UNIVERSITATEA BABEȘ-BOLYAI
ȘCOALA DOCTORALĂ: ȘTIINȚE ECONOMICE ȘI
GESTIUNEA AFACERILOR

TEZĂ DE DOCTORAT

– REZUMAT –

Rentabilitatea sistemelor (semi-)automatizate de
detectare a fraudei în asigurările auto din România

Conducător de doctorat:

Prof. univ. dr. Cristina CIUMAȘ

Doctorand:

Botond **BENEDEK**

2020

ABSTRACT

Începând din anii 1990, diverse studii au încercat să definească și să cuantifice fraudă în domeniul asigurărilor auto, să identifice predictorii cheie pentru cazurile de fraudă și să propună abordări privind metodele de detectare a fraudei. Cu toate acestea, până astăzi, nu există o listă general acceptată ai predictorilor de fraudă sau o metodă de detectare generală. Așadar, acest studiu are un triplu scop. În primul rând, să ofere o revizuire comprehensivă și sistematică a literaturii de specialitate pe tematica subiectului analizat de noi. După prezentarea detaliată a fenomenului de detectare a fraudei auto și a literaturii de specialitate (inclusiv dificultățile măsurării performanței metodelor de detectare a fraudei în cazul seturilor de date dezechilibrate), propunem o metodă îmbunătățită bazată pe economisirile generate. În final, comparăm ”abilitatea,, de economisire a 77 (respectiv 40) metode de detectare a fraudei folosind datele reale furnizate de către societățile de asigurare care au în portofoliu și subscriu asigurări auto (asigurări auto-casco și asigurarea obligatorie de răspundere civilă auto) reprezentative pentru piața de asigurare din România. Rezultatele analizei noastre arată clar, că tehnicile bazate pe *data mining* devin în ce în ce mai populare, dar evidențiază lipsa abordărilor sensibile la costurile implicate și lipsa seturilor de date relevante pentru cercetări academice și semnificative ca utilitate pentru asigurători. Considerând contextul românesc actual, rezultatele arată că doar un procentaj mic din metodele curente de detectare a fraudei auto pot fi aplicate într-un mod profitabil.

Cuvinte cheie: *asigurări auto; fraudă în asigurări; detectarea fraudei; revizuire a literaturii de specialitate; data mining; luarea deciziilor privind utilitatea metodelor în funcție de sensibilitatea la cost.*

CUPRINS

Abstract.....	5
Mulțumire.....	6
Abrevieri.....	9
Lista tabelor.....	10
Lista figurilor	11
INTRODUCERE.....	12
CAPITOLUL 1 – Industria Asigurării Auto.....	15
1.1 Apariția asigurării auto.....	15
1.2 Tipurile polițelor de asigurări auto și acoperiri disponibile	16
1.2.1 Tipurile generale a polițelor de asigurări auto	17
1.2.2 Acoperirile disponibile pe piețele dezvoltate.....	17
1.3 Răspunderile și legislația.....	19
1.3.1 Răspunderi legale generale în asigurarea auto.....	19
1.3.2 Obligații după un accident – coordonate generale.....	20
1.3.3 Legislația specifică la nivelul Uniunii Europene	21
1.3.4 Legislația în domeniul asigurării auto în România	22
1.4 Piața de asigurări auto	23
1.4.1 Piața de asigurări auto în SUA, Canada, Japonia și China	23
1.4.2 Piața europeană de asigurări auto	26
1.4.3 Piața română de asigurări auto.....	29
CAPITOLUL 2 – Frauda financiară, frauda în asigurări și frauda în asigurările auto 31	
2.1 Frauda financiară.....	31
2.1.1 Prezentarea tipurilor de fraudă financiară.....	32
2.1.2 Clasificarea metodelor și tehnicilor de data mining folosite pentru a preveni frauda financiară.....	34
2.2 Frauda în asigurări.....	36
2.2.1 Scheme de fraudă în asigurări.....	38
2.2.2 Costul fraudei în asigurări.....	41
2.2.3 Consecințele fraudei în asigurări ilustrate cu exemple	46
2.2.4 Combaterea fraudei în asigurări.....	48
2.2.5 Prezentare sintetică a unităților speciale de investigare.....	54

2.3	Frauda în asigurările auto.....	59
2.3.1	Tipurile fraudei de asigurări auto.....	59
2.3.2	Frecvența și costurile fraudei de asigurări auto	61
CAPITOLUL 3 – Detectarea Fraudei în Asigurările Auto.....		63
3.1	Abordarea fraudei auto în literatura de specialitate	63
3.2	Procesul general de detectare a fraudei în asigurările auto	71
3.3	Metode bazate pe data mining folosite în detectarea fraudei în asigurările auto	72
3.4	Predictorii fraudei în asigurările auto.....	83
3.3.2	Predictorii fraudei auto în literatura de specialitate	83
3.4.2	Clasamentul predictorilor de fraudă privind daunele materiale.....	88
3.4.3	Clasamentul agregat al predictorilor de fraudă privind daunele materiale	94
3.5	Seturile de date dezechilibrate și măsurarea performanței metodelor de detectare a fraudei în asigurările auto.....	97
3.5.1	Problematika seturilor de date dezechilibrate	98
3.5.2	Măsurarea performanței de clasificare a metodelor de detectare a fraudei în asigurările auto	102
3.6	Prezentarea sistematică a cercetărilor studiate	105
CAPITOLUL 4 – Rentabilitatea metodelor de detectare a fraudei în asigurările auto din România.....		114
4.1	Metoda proprie propusă de calculare a abilității de economisire.....	114
4.1.1	Metode de calculare a abilității de economisire în literatura de specialitate ...	116
4.1.2	Metoda personală de calculare a abilității de economisire	119
4.2	Studiu de caz pe un eșantion român al societăților de asigurare auto	120
4.3	Abilitatea de economisire a metodelor de detectare a fraudei auto ilustrată de către diagramele de tip <i>heatmaps</i>	124
CONCLUZII		127
Bibliografie		133
Referințe electronice		145
Anexe.....		146

Abrevieri

ALFA,	Agenție Împotriva Fraudei de Asigurări
CCI,	Asigurare Auto Completă (CASCO)
DT,	Arbore Decizional
FBI,	Biroul Federal de Investigații
FCM,	Fuzzy C-Means Clustering
FFD,	Detectarea Fraudei Financiare
FN,	False Negative
FP,	False Positive
FSA,	Autoritatea de Supraveghere Financiară
GA,	Algorithm Genetic
GMDH,	Metodă de Grup pentru Gestionarea Datelor
IAATI,	Asociația Internațională a Investigatorilor de Furt Auto
IFB,	Biroul de Investigare a Fraudelor în Asigurări
IFED,	Departamentul Impunerii Fraudei în Asigurări
IPD,	Asociația de Asigurări și Pensie din Danemarca
KNN,	K-Nearest Neighbors
MTPL,	Asigurare Obligatorie de Răspundere Civilă Auto
MLP,	Perceptron cu Multe Nivele
RF,	Câmp aleatoriu
SIU,	Unitatea Specială de Investigații
SVM,	Mașină Vectori Suport
TP,	True Positive
TN,	True Negative
UNESPS,	Uniunea Companiilor de Asigurări și Reasigurări din Spania
m,	milion
bn,	miliard

Lista Tabelelor

Tabelul 1: Metode bazate pe data minig (cu avantajele și limitele acestora) folosite în detectarea fraudei în asigurările auto	81
Tabelul 2: Predictorii de fraudă privind daunele materiale, frecvența de apariție și semnificația acestora	85
Tabelul 3: Predictorii de fraudă legați de vătămări corporale, frecvența de apariție și semnificația acestora.....	87
Tabelul 4: Clasamentul predictorilor de fraudă privind daunele materiale	89
Tabelul 5: Clasamentul agregat al predictorilor de fraudă privind vătămările corporale	94
Tabelul 6: Coeficienții de corelație de rang Kendall cu p-valorile acestora.....	97
Tabelul 7: Matricea de distribuție a unui clasificator binar	102
Tabelul 8: Cele mai folosite matrici de performanță derivate din matricea de bază	104
Tabelul 9: Precizia, sensibilitatea și specificitatea metodelor aplicate	105
Tabelul 10: Clasificarea metodelor de detectare a fraudei în asigurările auto pe baza algoritmului de bază	108
Tabelul 11: Seturile de date folosite în cercetările de detectare a fraudei în asigurările auto	109
Tabelul 12: Distribuția anuală a cercetărilor privind detectarea fraudei în asigurările auto..	111
Tabelul 13: Distribuția articolelor privind detectarea fraudei în asigurările auto în revistele de specialitate	112
Tabelul 14: Matrice de distribuție a unui clasificator binar	114
Tabelul 15: Matrice de cost a unei metode de detectare a fraudei auto.....	118
Tabelul 16: Potențialul de economisire a metodelor de detectare a fraudei auto	123
Tabelul 17: Potențialul de economisire a metodelor de detectare a fraudei auto în cazul a două companii românești de asigurare	124

Lista Figurilor

Figura 1: Procesul general de detectare a fraudei în asigurările auto	71
Figura 2: Reprezentare grafică a unei rețele bayesiană indicând o relație obișnuită între ipotezele H_0 și H_1	73
Figura 3: Rețea neuronală simplă cu trei variabile de intrare, două produse, și un singur strat ascuns	74
Figura 4: Exemplu bidimensional arătând marginea de separare între vectori suportați și hiperplane	76
Figura 5: Procesul standard al unui algoritm genetic.....	77
Figura 6: Arbore de decizie binară unde variabilele de intrare X_i vor fi o listă de predicate de intrare și produsele Y_i vor clasifica rezultatele într-una dintre mai multe clase	78

INTRODUCERE

Frauda în asigurările auto este o problemă care are consecințe importante atât în industria asigurărilor cât și în viața cotidiană. Frauda poate să producă neîncredere în industrie, să destabilizeze economii și să afecteze costul vieții oamenilor. Sondaje și studii începând din anii 1990 încearcă să cuantifice dimensiunile acestei activități, să identifice predictorii de fraudă importanți și să ofere metode și modele care poate să indice și să prognozeze solicitările posibil frauduloase în industria asigurării auto. Cu toate acestea, până astăzi, nu există o listă acceptată universal de predictorii de fraudă sau un model general. În cele mai multe țări, dimensiunile reale ale fraudei sunt necunoscute. Mai mult, cei mai mulți autori în literatura de specialitate au folosit seturile de date proprii cu diferiți predictorii de fraudă și le-au aplicat pe un model dezvoltat independent, a cărui performanță a fost măsurat de o metrică arbitrară.

Relevanța și importanța subiectului este evidențiată de multiplele rapoarte publicate pe subiect, care afirmă, că frauda în asigurări afectează 10-20% din totalul contractelor încheiate chiar și în cele mai dezvoltate țări. Așadar, frauda în asigurări este a doua cea mai răspândită fraudă financiară, după frauda fiscală. Un raport FBI afirmă că o familie obișnuită în USA are \$400-\$700 cheltuieli în plus în fiecare an datorită fraudei de asigurări (FBI, 2011). Dacă ne uităm doar la asigurarea auto, în SUA și Europa de Vest, 7-10% din polițele sunt afectate, iar în Europa Centrală și de Est, această cifră ajunge la 10-20%. În România, nu există statistici oficiale privind frauda în asigurarea auto. Totodată, Thomas Brinkmann, managerul teritorial a Friss-ului, a declarat la Forumul Internațional Asigurări-Reasigurări în 2017, că cel puțin 15% din asigurările auto sunt afectate în România. Această afirmație pare a fi susținută și de un raport ASF din 2016, care estimează cifra de 16%. Din punct de vedere financiar, aceste cifre reprezintă o fraudă de 1.7-1.8 miliarde de RON pe an, care este echivalentul a 400 milioane de euro.

Așadar, acest studiu al nostru are un triplu rol. În primul rând, la fel ca (Ngai, et al., 2011), (West & Bhattacharya, 2016), și (Abdallah, et al., 2016), acest studiu încearcă să prezinte o revizuire sistematică și comprehensivă a literaturii academice, concentrându-se strict pe detectarea fraudei în asigurările auto (contrar cu publicațiile menționate anterior, care au accentuat detectarea fraudei financiare, în general). 44 articole din reviste de specialitate și 8 publicații din volume ale unor conferințe legate de subiect, publicate între anii 1990 și 2019,

vor fi analizate, studiu bazat pe aspecte cheie cum ar fi: setul de date, predictorii de fraudă, algoritmul de detectare respectiv performanța metodelor de detectare.

După prezentarea detaliată a literaturii de specialitate privind fraudă în asigurările auto inclusiv literatura despre luarea deciziilor de utilizare a metodelor în funcție de sensibilitatea la cost (*cost-sensitive decision making*) și dificultățile de măsurare a performanței în cazul seturilor de date dezechilibrate, cum ar fi seturile de date pentru fraudă în asigurările auto), propunem o metodă îmbunătățită de calculare a economisirilor (care ia în calcul toate costurile aferente procesului de verificare a dosarelor de despăgubire) pentru a oferi o metrică ușor de folosit pentru a compara abordări diferite din literatura de specialitate.

În final, comparăm abilitatea de economisire a 77 (respectiv 40) de metode de detectare a fraudelor în asigurările auto folosind date din România. În acest mod, vom putea să răspundem la întrebarea principală a cercetării, mai exact, pot fi aplicate metodele de detectare a fraudei în asigurarea auto într-un mod profitabil?

Desigur, pe durata analizării literaturii de specialitate și a constatării golurilor, a conturării problematicii evidențiate mai devreme, apar mai multe întrebări referitoare la subiect, cum ar fi: *care sunt cei mai importanți predictorii de fraudă? care este cea mai eficientă metodă de detectare a fraudei în condiții comerciale și economice particulare?* Am putea să adăugăm aici și întrebări în legătură cu literatura de specialitate, precum: *în ce măsură sunt seturile de date reprezentative, având în vedere actualul mediu comercial?*

Pentru a obține un răspuns la aceste întrebări secundare, am completat studiul principal cu mai multe subcapitole, care nu sunt strict, imperios necesare pentru a răspunde la întrebarea principală a cercetării, însă ei au legătură cu subiectul tratat și oferă ajutor pentru a găsi aceste răspunsuri, spre exemplu, capitolul în care identificăm și discutăm predictorii de fraudă și construim un clasament agregat bazat pe mai multe perspective sau capitolul în care prezentăm heatmap-urile care ajută să înțelegem performanța unor metode de detectarea a fraudei în condiții comerciale și economice diferite, astfel ușurând munca viitoare a oricărei societăți de asigurare sau cercetător.

Pe lângă subiectele enumerate, cercetarea încearcă să ofere cititorului o prezentare generală despre fraudă financiară și fraudă în asigurări, categoriile acestora, legislațiile aferente și, nu în ultimul rând, implicațiile economice ale acestora.

Restul studiului este organizat în felul următor: în primul capitol, după o scurtă prezentare a apariției asigurării auto, legile actuale care reglementează asigurarea auto, subliniind legile din Uniunea Europeană și România sunt analizate diferite tipuri de asigurări auto și acoperiri disponibile. Acest capitol oferă informații și despre trendurile actuale ale pieței de asigurări auto.

În capitolul al doilea, diferite scheme de fraudă în asigurări vor fi prezentate, alături de diferite metode folosite pentru prevenție și detectare. Capitolul acesta dedică o atenție specială fraudei în asigurări. În afară de acestea, tehnicile, pe care departamentele de antifraudă le folosesc pentru a combate fraudă în asigurări, vor fi prezentate. În ultima parte a capitolului, vom oferi o analiză detaliată a schemelor de fraudă în asigurările auto, cu mărimea și costurile acestora în anumite țări.

În al treilea capitol, ne vom concentra în special pe fraudă în asigurările auto și metodele de detectare a acesteia. În afară de prezentarea literaturii de specialitate, vom analiza cei mai importanți predictorii de fraudă și cele mai frecvent utilizate metode de detectare a fraudei, bazate pe data mining sau inteligență artificială, accentuând avantajele și dezavantajele acestora. În plus, vom revizui seturile de date prezente în literatura de specialitate și metricile de măsurare specifice folosite pentru măsurarea performanței.

În al patrulea capitol, vom analiza abilitatea de economisire a metodelor studiate de detectare a fraudei în contextul actual românesc. Având în vedere performanța metodelor de detectare a fraudei în funcție de contextul economic și de alte condiții industriale posibil diferite, construim heatmap-uri pentru a simula performanța tuturor metodelor de detectare.

Ultimul capitol rezumă rezultatele și prezintă limitările studiului.

SINTEZA: CAPITOLUL 1 – INDUSTRIA DE ASIGURĂRI AUTO

Când automobilul a fost inventat – la sfârșitul secolului al 19-lea – nevoia asigurării nu a apărut instant. În câteva țări, precum în Marea Britanie, a fost în vigoare o lege de siguranță: o persoană trebuia să se plimbe în fața unei mașini în deplasare cu un steag roșu pentru a avertiza șoferul să nu lovească pietonii. În 1896 reglementarea a fost revocată și industria auto britanică a început să se dezvolte. În 1989 a fost emisă prima poliță de asigurare auto, chiar dacă asigurarea comprehensivă a apărut doar câțiva ani mai târziu (Benetton, 1994).

De atunci, industria de asigurări auto a devenit una dintre cele mai mari industrii. Spre exemplu, în 2018, numărul vehiculelor înregistrate în SUA era mai mult de 273 milioane, și 215 milioane în Europa. În plus, doar în Statele Unite, prima de asigurare aferentă asiguraților privați a depășit 148 miliarde dolari în 2019.

Datorită dezvoltării accelerate menționate a industriei de asigurări auto, în primul capitol prezentăm câțiva indicatori-cheie ai pieței de asigurări auto, cum ar fi numărul polițelor subscrise, accidentelor, persoanelor decedate sau prima medie de asigurare, la nivelul pieței europene și românești.

În afară de prezentarea cifrelor cheie a piețelor de asigurări, în capitolul întâi sunt prezentate și diferite tipuri de polițe de asigurare auto și acoperirile disponibile. Bazat pe (Sims, 2019), următoarele acoperiri sunt cele mai folosite pentru a proteja proprietarii auto asigurați, părți terțe și bunurile acestora: acoperirea cheltuielilor medicale, acoperire a răspunderii, acoperire a daunei materiale și acoperire a șoferilor neasigurați. Fiecare dintre acestea alături de specificațiile lor, sunt prezentate în primul capitol.

După prezentarea diferitelor tipuri de polițe de asigurări auto și specificațiile acestora, este esențial să clarificăm responsabilitățile acoperite de contractele de asigurare și reglementările legale asociate, cel puțin în cazul Uniunii Europene și România. Evident, reglementările înșăși sau metodele folosite pentru a determina compensații, diferă între țări, dar pot fi găsite câteva principii comune. Acest lucru este în special adevărat în cazul pieței europene, unde reglementările sunt armonizate la nivelul Uniunii Europene.

SINTEZA: CAPITOLUL 2 – FRAUDĂ FINANCIARĂ, FRAUDA ÎN ASIGURĂRI ȘI FRAUDA ÎN ASIGURĂRILE AUTO

Înainte de toate, capitolul al doilea a tezei oferă o definiție clară și detaliată a conceptelor precum: *frauda*, *dosar fraudulos*, *frauda în asigurări* și *detectarea fraudei*. După acestea, în prima parte a capitolului al doilea, sunt prezentate diferite scheme de fraudă financiară și în asigurări, împreună cu diferite metode de *data mining* folosite pentru prevenirea și detectarea fraudei. Bazată pe Raportul de Crime Financiare publicată de FBI, fraudă financiară poate fi împărțită în 7 categorii diferite: *fraudă de card de credit*, *fraudă de titluri* și de mărfuri, *fraudă de situații financiare*, *fraudă în asigurări*, *fraudă ipotecară* și *spălarea banilor* (FBI, 2011).

În partea a doua parte a acestui capitol, analizăm metodele de detectare și prevenire, care sunt cele mai folosite de instituțiile financiare. (Ngai, 2011) împarte aceste metode în următoarele 6 categorii: clasificare, grupare, predicție, detectarea valorilor extreme, regresare și vizualizare. Principiile generale de operare a acestor șase categorii, avantajele și dezavantajele acestora sunt prezentate în această secțiune.

Focusând pe scopurile tezei curente, am dedicat atenție specială fraudei în asigurări, prezentând modul de funcționare, caracteristicile acesteia, statisticile și trendurile aferente. În plus, prezentăm pe scurt și tehnicile care sunt folosite de departamentele anti-fraudă pentru a combate fraudă în asigurări.

În ultima parte a capitolului, oferim o prezentare detaliată a fraudei în asigurările auto, împreună cu dimensiunile și costurile acestor activități în diferite țări. Bazat pe literatura de specialitate, pe sondaje industriale și pe rapoarte ale diferitor corpuri investigative, am identificat 14 diferite tipuri de fraudă în asigurările auto. Acestea sunt următoarele: 1) abandonarea autovehiculului și reclamarea furtului; 2) subscrierea poliței de asigurare după producerea accidentului; 3) repararea vehiculului și solicitarea acoperirii prin poliță fără ca acest lucru să fie produs de un eveniment asigurat, 4) asigurarea autovehiculelor provenite din contrabandă; 5) vehicule "fantomă", 6) accidente aranjate, 7) supraevaluarea daunelor, 8) schimbarea numărului de identificare al vehiculului, 9) fraudă cu mașină închiriată, 10) scheme cu alte bunuri implicate, 11) inventar "umflat", 12) furt simulat, 13) furt supraestimat, 14) incendierea autovehiculului pentru "profit". În afară de identificarea tipurilor de fraudă în asigurările auto, frecvența și costurile fraudelor auto sunt prezentate detaliat.

SINTEZA: CAPITOLUL 3 – DETECTAREA FRAUDEI ÎN ASIGURĂRILE AUTO

În capitolul al treilea analizăm fraudă în asigurările auto și metodele de detectare a acesteia. Înainte de toate, prezentăm literatura de specialitate. Am reușit să identificăm 52 articole științifice și publicații în volum de conferință (după 1990) care sunt indexate de către Web of Science și sunt legate de domeniul fraudei în asigurările auto. O parte din aceste studii au abordat domeniul fraudei de asigurări auto teoretic (spre exemplu (Weisberg & Derrig, 1991); (Derrig & Ostaszewski, 1995); (Derrig, 2002); (Tennyson & Salsas-Forn, 2002); (Warren & Schweitzer, 2018)), iar ceilalți oferă un sistem/model/abordare specific de detectare a fraudei (precum (Belhadji & Dionne, 1998); (Artís, et al., 1999); (Artís, et al., 2002); (Viaene, et al., 2002); (Viaene, et al., 2005); (Viaene, et al., 2007); (Pinquet, et al., 2007); (Phua, et al., 2004); (Bermúdez, et al., 2008); (Wilson, 2009); (Sundarkumar, et al., 2015); (Sundarkumar & Ravi, 2015); (Hassan & Abraham, 2016); (Subudhi & Panigrahi, 2017); (Wang & Xu, 2018); (Majhi, et al., 2019); (Zelenkov, 2019)).

După prezentarea literaturii de specialitate, descriem detaliat procesul de detectare a fraudei. În acest stadiu, este necesar să menționăm, că nici una dintre tehnologiile anti-fraudă existente nu pot demonstra fraudă în sine. Investigații de nivel înalt sunt cruciale pentru combaterea fraudei și în prezent tehnologia nu poate substitui un investigator experimentat și priceput. Totuși, tehnologia oferă un ajutor imens în procedura așa-zisă de depistarea precoce a dosarelor frauduloase, care pot fi făcută și cu un model (semi-)automat care efectuează depistarea fraudei bazată pe predictorii. Acest model poate implementa tehnologii diferite, începând cu cele mai simple sisteme, până la cele mai complexe metode. Așadar, partea următoare a capitolului al treilea oferă o prezentare detaliată a metodelor existente, alături de avantajele și dezavantajele acestora.

În timpul dezvoltării oricărui sistem de detectare a fraudei (semi-)automată, este extrem de important să definim corect predictorii potențiali de fraudă. Din cauza acestora, în următoarea parte a capitolului prezentăm cei mai importanți predictorii de fraudă legate de vătămarea corporală și daunele materiale. Totuși, discutăm doar predictorii privind daunele materiale în detalii, deoarece acest tip de predictor este accesabil în timpul procedurii depistării precoce. Predictorii legați de vătămările corporale sau spitalizare și recuperare nu sunt disponibili în acest stadiu. Privind predictorii de daune materiale, am compilat o listă de 92 "indicatori". În

timpul creării acestei liste, am considerat atât literatura de specialitate, cât și opiniile experților industriei din România. Bazat pe clasamentele create din diferite puncte de vedere, am construit un clasament agregat, care poate fi folosit ca punct de plecare pentru cercetări viitoare și pentru lucrarea oricărui cercetător sau expert al industriei de asigurare.

După prezentarea diferitelor metode de detectare a fraudei și predictorilor de fraudă, urmează descrierea măsurării performanței în cazul metodelor de detectare a fraudei de asigurări auto. Măsurarea performanței a metodelor de detectare a fraudei (inclusiv, dar nu limitat la fraude de asigurări auto) este o problemă de clasificare binară. Totuși, măsurarea performanței a metodelor de detectare a fraudei este specială, din cauza problemei generate de setul de date dezechilibrate și impactul acestuia asupra măsurării performanței trebuie luat în considerare.

În final, pentru a rezuma informațiile din cele 52 studii despre detectarea fraudei în asigurările auto, cu respect la provocările prezentate mai sus, în ultima parte a capitolului prezentăm diferite tabele cu clasificări și sintetizări. Cu această compartimentare putem prezenta trendurile în metodele de cercetare, accentuând cele care au avut succes și factorii posibili nediscuțați până acum. Mai mult, performanța metodelor de detectare, seturile de date, cele mai importante cercetări și reviste devin ușor de revizuit.

SINTEZA: CAPITOLUL 4 – RENTABILITATEA METODELOR DE DETECTARE A FRAUDEI ÎN ASIGURĂRILE AUTO DIN ROMÂNIA

În al patrulea capitol, bazat pe provocările legate de detectarea fraudei în asigurările auto prezentate anterior, analizăm rentabilitatea fiecărei metode de detectare prezentate. Este evident, că din punct de vedere corporativ este necesar să măsurăm beneficiul (reducerea costului) a oricărei metode de detectare. Pentru a calcula acest beneficiu generat de folosirea unei anumite metode de detectare, punctul nostru de plecare a fost dat de (Phua et al., 2004) și (Viaene, et al., 2007).

Este important să remarcăm, că în ambele cazuri autorii își fundamentează demersul plecând de la presupunerea că, *cazurile negativ reale* (când un dosar legal este marcat legal) nu au extra cheltuieli pentru companiile de asigurări. Totuși, după interviuarea a mai multor experți i români ai industriei de asigurare și manageri ai departamentelor de anti-fraudă, am ajuns la concluzia, că în realitate aceste *cazuri negativ reale* au costuri suplimentare. Cheltuielile acestea pot fi ori salariul unui angajat care îndeplinește sarcini specifice legate de procesul de verificare a legitimității a unui dosar, sau pur și simplu costul unui serviciu folosit. Chiar dacă acești cheltuieli sunt doar de câțiva euro, considerând numărul semnificativ al dosarelor, se pot aduna la o valoare imensă.

Similar cu *cazurile negativ reale*, calcularea costului *cazurilor pozitiv reale* (când un dosar fraudulos este marcat fraudulos) este diferit în realitate de cum este descris în literatura de specialitate. Cauza fiind, că în literatură în cazul unei dosar pozitiv reale, compania de asigurări nu plătește asiguratul și singurul cost este cel al investigației. Cu toate acestea, în realitatea industriei asigurărilor, situația este diferită. Multe studii, inclusiv (Derrig & Ostaszewski, 1995), (Weisberg & Derrig, 1998), au dovedit că fraudă în asigurările auto în mare parte este reprezentată de contracte supraevaluate (build-up). Când o fraudă de tip build-up este detectată, societatea de asigurare plătește o despăgubire redusă. Interviuații noștri au confirmat că în realitate, cazurile în care societatea refuză să plătească despăgubirea, sunt foarte rare (în general asiguratorul oferă o sumă mai mică, pe care o consideră corectă). Această decizie are multe explicații, date fiind procesul juridic lung sau marketingul negativ.

Datorită problemelor prezentate anterior, în capitolul al patrulea am propus o nouă metodă de calculare a costurilor, care ia în considerare toate cheltuielile care apar de-a lungul procesului de detectare a fraudei. Profitabilitatea metodelor de detectare a fraudei prezentate

anterior în cele 52 de articole științifice și publicații în volume de conferință, care sunt legate de domeniul fraudei auto, a fost determinată folosind metoda proprie de calculare a costului. Este important de observat, că o parte din aceste studii au abordat domeniul fraudei în asigurările auto teoretic și nu oferă un sistem specific de detectare a fraudei. Din această cauză, am putut identifica 77 metode de detectare a fraudei bazate pe data mining sau inteligență artificială. Cu toate acestea, cele 77 de metode de detectare conțin câteva abordări ale căror performanțe sub circumstanțe 100% reale este discutabilă¹, așadar am analizat și o listă compusă din doar 40 de modele (modele pentru care performanța metodei de detectare a fost testată folosind seturi de date reale, dezechilibrate). Tabelul 1 rezumă abilitatea de economisire a celor 77 (respectiv 40) modele în trei cazuri diferite.

Tabelul 1: Abilitatea de economisire a metodelor de detectare în cazul companiilor de asigurări auto din România.

	Cazul cel mai probabil		Cazul cel mai nefavorabil		Cazul cel mai favorabil	
	77 modele	40 modele	77 modele	40 modele	77 modele	40 modele
Procentul dosarelor frauduloase	10%	10%	5%	5%	20%	20%
Despăgubirea medie (euro/ dosar)	2420	2420	2420	2420	2420	2420
Costul mediu de investigație (euro/ investigație)	145	145	193	193	97	97
Costul mediu de procesare a unei dosar nefraudulos (euro/dosar)	12	12	12	12	12	12
Economisirea medie în cazul dosarelor frauduloase identificate (euro/ dosar)	485	485	315	315	1213	1213
Numărul modelelor neprofitabile	41	28	70	39	0	0
Numărul modelelor cu economisire						
>25% din maximul posibil	26	5	5	1	71	35
>50% din maximul posibil	17	1	3	1	63	32
>75% din maximul posibil	11	1	1	1	34	9
>90% din maximul posibil	0	0	0	0	7	0

Sursă: prelucrări personale

Metoda de calculare a economisirii propuse de noi a fost aplicată și în cazul a două companii de asigurări pentru a verifica profitabilitatea a 77 (respectiv a 40) modele (vezi Tabelul 2).

¹ Spre exemplu, (Artís, et al., 1999) și (Artís, et al., 2002) au folosit un set de date reale în timpul studiului, dar eșantionul nu a fost selectat aleatoriu, deoarece a fost compus din 997 dosare frauduloase și 998 dosare legitime. (Wilson, 2009) a folosit 49 dosare frauduloase și 49 dosare legitime. (Wang & Xu, 2018) au lucrat cu 1660 dosare frauduloase și 1660 dosare legitime. Performanța reală a metodelor menționate pe seturi de date dezechilibrate este incertă.

Dosarele de RCA și de CASCO au fost analizate separat pentru ambele societăți de asigurare.

Tabelul 2 rezumă rezultatele.

Tabelul 2: Abilitatea de economisire a metodelor de detectare în cazul a două companii românești de asigurare.

	Compania A		Compania B	
	RCA	CASCO	RCA	CASCO
Procentul dosarelor frauduloase	10%	3%	12%	8%
Despăgubirea medie (euro/ dosar)	1250	870	2360	1160
Costul mediu de investigație (euro/ investigație)	205	97	160	160
Costul mediu de procesare a unei dosar nefraudulos (euro/dosar)	12	12	0	0
Economisirea medie în cazul dosarelor frauduloase identificate (euro/ dosar)	160	242	675	338
Numărul modelelor neprofitabile	40	33	22	36
Numărul modelelor cu economisire				
>25% din maximul posibil	0	4	7	1
>50% din maximul posibil	0	1	2	1
>75% din maximul posibil	0	1	1	1
>90% din maximul posibil	0	0	0	0

Sursă: prelucrări personale

Rezultatele în ambele cazuri arată clar că doar un procentaj mic din modelele de detectare a fraudei disponibile momentan poate fi aplicate într-un mod profitabil.

Bineînțeles, observând rezultatele anterioare, apar diferite întrebări, precum: ce s-ar întâmpla dacă am folosi diferite inputuri în cazul determinărilor implicate de utilizarea acestei metode proprii? sau metodele de detectare a fraudei sunt neprofitabile și în cazul altor companii? Pentru a răspunde la aceste întrebări, abilitatea de economisire a celor 40 de metode a fost calculat în 3534 situații diferite, acoperind fiecare combinație de inputuri posibile. Rezultatele scenariilor sunt ilustrate cu ajutorul diagramelor de tip heatmaps (hărților de căldură).

CONCLUZII

O parte esențială a oricărui studiu este dată de construirea unui cadru de clasificare competent și unei colecții a literaturii de specialitate pe tema supusă analizei. Cercetarea detectării fraudei în asigurările auto nu este o excepție. Deși importanța metodelor de *data mining* în detectarea fraudei auto este recunoscută, cadrul menționat specific lipsește. Din această cauză, în primul rând, am realizat o revizuire a articolelor științifice și am conturat o bibliografie de referință și un cadru de clasificare extensivă pentru utilizarea metodelor de data mining și machine learning în detectarea fraudei în domeniul asigurărilor auto. În acest sens, scopul nostru a fost să oferim informații valoroase atât pentru cercetători cât și pentru practicieni în scopul utilizării eficiente a acestor tehnici de detectare a fraudei în sfera asigurărilor auto și să realizăm o revizuire organizată a publicațiilor semnificative legate de detectarea fraudei auto.

Referitor la partea mai sus menționată a studiului (revizuire comprehensivă a literaturii de specialitate), am dori să accentuăm câteva dificultăți generale pe care le-am întâlnit și sunt esențiale pentru a înțelege literatura aferentă. În primul rând, studiile au arătat că nu există o definiție general acceptată a fraudelor în asigurările auto. Definiția cea mai frecvent utilizată la nivel mondial este cea din Massachusetts Regulation (211 CMR 93.03) care definește dosarele frauduloase ca dosare depuse cu intenția de a primi o compensare mai mare de la asigurator decât suma, la care solicitantul are dreptul potrivit poliței de asigurare încheiate, inclusiv cererile pentru (i) pierderile inexistente; (ii) sumele care depășesc pierderile reale; sau (iii) incidente care sunt organizate de solicitant pentru a obține o compensare de la asigurator (Massachusetts Regulation, 1993). Cu toate acestea, această definiție nu conține celelalte tipuri de fraudă, precum interpretarea greșită a condițiilor de asigurare, denaturarea sau neglijența persoanei asigurate. Cum este prezentat și în (Wilson, 2009), denaturarea așa-zisă nu este atât de intenționată. Fraudatorii pot să își justifice astfel, că acțiunile lor sunt în regulă. Un exemplu ar fi un părinte a cărui copil recent a primit permisul de conducere și care nu anunță asiguratorul că copilul folosește mașina pur și simplu din cauza faptului că tarifele de asigurare a familiei vor crește foarte mult în acest caz (Wilson, 2009). Asemănător, neglijența nu presupune neapărat o încercare intenționată să producă o pagubă. Cu toate acestea, o persoană asigurată se angajează în acțiuni în care în mod normal nu s-ar fi angajat, dacă rezultatul posibil advers ar fi ”povara” personală. Un exemplu ar fi când oamenii au încheiate asigurări CASCO și

devin mai puțin atenți din cauza acestor subscrieri. Lângă excepțiile menționate mai sus, este importat să notăm faptul că numeroase țări au legislații diferite care duc la diverse produse de asigurare. Produse diverse generează comportamente diferite ale asiguraților și din cauza acestor unele tipuri de fraudă auto devin mai importante în anumite țări. Un exemplu ar fi fraudarea asigurărilor legate de vătămări corporale, care sunt importante în SUA, dar nu sunt atât de relevante în context european (din cauza asigurării obligatorii de sănătate) (Artís, et al., 1999). În concluzie, lipsa unei definiții precise a fraudei în asigurările auto poate fi explicată de varietatea produselor de asigurare auto, diversitatea legislației la nivel de țări și comportamentul diferit al persoanelor.

O altă problemă generală în detectarea fraudei auto este lipsa datelor credibile. Cum este prezentat și în (Brockett, et al., 1998) fiabilitatea fiecărei variabile dependente nu poate fi verificată în lumea reală din cauza limitărilor generate de disponibilitatea datelor. Spre deosebire de alte probleme de detectare a fraudei, precum fraudă de carduri de credit, cele mai multe dosare de despăgubire cu componente de vătămări corporale, până la urmă nu pot fi verificate, ori procesul de verificare nu este rentabil sau este imposibil de realizat, cu atât mai mult în cazul unei fraude fără o hotărâre judecătorească disponibilă. Totodată, societățile de asigurare au tendința să nu rezolve cererile în acest fel, deoarece sunt riscante și scumpe (Brockett, et al., 1998). Prin urmare, seturile de date (nu doar dosarele cu componente de vătămare corporală, ci toate dosarele de asigurări auto) folosite în detectarea fraudei, în cele mai multe cazuri conțin predictorii și clasificării subiective. Această problemă a fost prezentată detaliat de (Weisberg & Derrig, 1991) și de (Derrig & Ostaszewski, 1995). Ei au constatat că profesioniștii care procesează diferite dosare cu componente de vătămări corporale au avut percepții ambigue referitoare la dosarele cu componente de vătămări corporale care reprezintă fraudă (Weisberg & Derrig, 1991). În plus, (Derrig & Ostaszewski, 1995) au demonstrat și lipsa concordanței de opinie între experți, referitoare la dosarele frauduloase. Totuși, acești predictorii subiectivi și clasificării făcute de experții industriei sunt singurele surse de date disponibile. Continuând cu problematica generată de bazele de date, am dori să accentuăm că în literatură au fost realizate studii fundamentate doar pe câteva seturi de date. Din cele 14 seturi de date studiate în 4 cazuri, anul pentru care au fost colectate este necunoscut și doar un set de date conține dosare de daune auto din ultimii zece ani, subliniind lipsa datelor credibile în cercetările academice. În plus, cea mai mare parte a acestor seturi de date acoperă doar trei piețe, mai exact au fost colectate pe piețele din Spania, China și SUA. Cele mai multe seturi de

date conțin doar câteva mii de dosare, în timp ce numărul anul real de dosare de daune auto este mult mai mare, astfel reprezentativitatea bazelor de date este discutabilă.

După analizarea literaturii, o altă concluzie importantă a studiului este dată de lipsa abordărilor care iau în considerare costurile aplicării diverselor metode. Chiar dacă scopul procesului de detectare a fraudei auto este reducerea cheltuielilor companiilor de asigurare cauzate de acest fenomen, există doar câteva studii de acest tip. Cele mai multe studii se focusează pe minimizarea ratei de eroare (clasificare eronată) mai degrabă decât să se analizeze costurile totale sau rentabilitatea procesului de investigare. Doar (Phua, et al., 2004), (Viaene, et al., 2007) și (Zelenkov, 2019) au propus, astfel de abordări.

După o revizuire comprehensivă a literaturii de specialitate, alături de diferite tipuri de fraudă specifice domeniului asigurărilor auto, în acest studiu am prezentat un clasament agregat a celor mai importanți predictorii ai fraudei auto. În timp ce cream clasamentul agregat, am ținut cont atât de predictorii prezentați în literatură cât și de predictorii recomandați de liderii industriei asigurărilor auto, identificați pe parcursul derulării interviurilor noastre. Așadar, clasamentul agregat poate servi ca un punct de plecare util pentru orice viitor cercetător sau asigurător, care este interesat de orice parte a detectării fraudei auto.

Ultima parte a studiului este legată de măsurarea performanței metodelor de detectare a fraudei în cazul seturilor de date dezechilibrate și abilitatea de economisire a acestora. Studiul arată că indicatorii de performanță utilizați în prezent ori nu pot fi folosiți pe seturi de date dezechilibrate (cel puțin nu în modurile prezentate în literatura de specialitate) ori comparația metodelor devine foarte dificilă, deoarece performanța unei metode este imposibil de cuantificat cu ajutorul unui singur indicator. În plus, acești indicatori de performanță nu iau în considerare costurile care apar de-a lungul procesului de detectare a fraudei. Prin urmare, studiul actual oferă o evaluare a performanței bazată pe cheltuieli în locul minimizării numărului de clasificări eronate. Deși există o abordare propusă de (Phua, et al. 2004) care ține cont de cheltuielile implicate, în timpul interviurilor realizate am notat următoarele: nu numai că procesul de detectare a fraudei este diferit ca abordare în literatura și respectiv practica industriei asigurărilor auto, dar și costurile aferente diferă. Luând în considerare procedurile reale și costurile implicate, studiul actual arată, că aplicarea celor mai multe modele recomandate de literatură, în contextul economic și comercial cel mai probabil, nu este rentabilă deloc.

Deoarece contextul economic și comercial se schimbă continuu, am creat mai multe scenarii cu inputuri (valori) medii de pe piață pentru a analiza abilitatea de economisire a metodelor. Mai mult, am folosit inputuri specifice oferite de către doi asigurători din România și am analizat abilitatea de economisire a modelelor în cazul acestor companii de asigurări examinând dosarele de CASCO și RCA separat. Rezultatele arată clar, că doar un procentaj redus din abordărilor de detectare a fraudei auto studiate pot fi aplicate într-un mod rentabil.

Pe lângă cele menționate mai sus, am creat ilustrări ce se doresc relevante cu ajutorul diagramelor de tip *heatmap* (hărți de căldură) pentru a compara abilitatea de economisire a metodelor de detectare a fraudei auto studiate, astfel oferind o posibilitate asigurătorilor și cercetătorilor interesați în domeniu pentru a alege simplu și eficient metoda cea mai optimă.

Provocările generate de subiectul detectării fraudei în domeniul asigurărilor auto și direcțiile viitoare de cercetare

Detectarea fraudei auto este un domeniu care se dezvoltă mereu și unde este optim și recomandat să fii cu un pas înaintea fraudatorilor. Este evident și faptul că multe aspecte ale detectării fraudei inteligente necesită cercetare. În secțiunea următoare vom arăta câteva din cele mai importante aspecte ridicate de problematica detectării fraudei auto și vom recomanda direcții pentru viitoare cercetări.

Considerând că aceasta este o problemă de clasificare, detectarea fraudei în asigurările auto suferă de aceleași probleme ca și domeniile similare. Selecția caracteristicilor are un impact imens asupra ratei de succes a oricărei metode de clasificare. Avantajul cel mai mare a tehnicilor bazate pe inteligență artificială și data mining este capacitatea acestora de a se adapta la specificul problemei. Doar câteva studii actuale au folosit orice fel de personalizare sau ajustare optimă pentru probleme specifice, chiar dacă ajustarea este un factor semnificativ privind performanța algoritmului. Spre exemplu, numărul nodurilor și ramurilor interne într-o rețea neuronală are un impact semnificativ asupra preciziei și performanței computaționale. De asemenea, funcția de tip *kernel* aleasă va modifica semnificativ rata de succes a unui algoritm de tip *suport vector machine* și criteriile precum funcția de tip *fitness*, tehnica transversală (crossover) și probabilitatea mutației vor afecta rezultatele a unui algoritm de tip programare genetică. Sunt necesare studii axate pe stimularea și personalizarea tehnicilor computaționale pentru a obține o perspectivă reală asupra abilității metodelor analizate.

Deoarece datele financiare sunt private, organizațiile rețin pentru sine informațiile legate de cazurile de fraudă. Aceste aspecte au afectat atât tipurile de fraudă analizate cât și seturile de date folosite. În publicațiile prezentate o parte semnificativă a metodelor de detectare a fraudei în asigurările auto au fost aplicate pe mai puțin de câteva mii de exemple, în general având un număr aproape identic de cazuri legale și respectiv frauduloase. Acest fenomen este total diferit de realitatea domeniului problematic, unde dosarele legitime depășesc cu mult cazurile frauduloase. Având puține exemple, este posibil să apară interpretări eronate asupra datelor care astfel, nu reprezintă precis, cazurile din lumea reală. Studii suplimentare sunt necesare cu exemple reale pentru a obține o reprezentare adecvată a performanței fiecărei metode.

Spre deosebire de cele mai multe probleme de clasificare, soluțiile de detectare a fraudei în asigurările auto trebuie să aibă capacitatea de a procesa noi tipuri de cazuri. Deoarece fraudatorii mereu își îmbunătățesc tehnicile, iar metodele menționate trebuie să se dezvolte la fel. Cercetătorii au analizat capacitatea de adaptare a metodelor, însă sunt necesare studii detaliate pentru a îmbunătăți cât mai mult aceste metode pentru a putea fi folosite în problemele de detectare a fraudei în lumea reală.

Cum am discutat anterior, costul fraudei pentru asigurători este imens la care se adaugă și costurile asociate: sistemele de detectare a fraudei au nevoie de întreținere și putere computațională, alături de experți pentru a le monitoriza constant și a analiza cazurile identificate ca potențial frauduloase. Costul clasificării eronate a unui dosar legitim este de obicei mult mai redus decât opusul acestuia. Doar câteva studii au fost publicate în legătură cu natura disproporționată a costurilor menționate. Considerând precizia, sensibilitatea și specificitatea fiecărei metode de detectare a fraudei, scopul principal ar trebui să fie atingerea unei aprofundări optime a utilității fiecărei metode, avându-se în vedere costul cel mai mic implicat. Studii orientate în mod explicit asupra găsirii acestui echilibru ar crește enorm valoarea practică a detectării fraudei în asigurările auto.

În legătură cu abilitatea de economisire a metodelor de detectare a fraudei, ar fi util să fie examinată combinația metodei noastre cu abordarea recomandată de (Phua, et al., 2004), prelucrând dosarele de daună totală (când suma compensației, despăgubirii este egală cu valoarea vehiculului asigurat) și daună parțială, evaluată pe bază de deviz, separat. Bineînțeles, o bază de date reale disponibilă, care separă aceste două tipuri de dosare este necesară pentru finalizarea demersului propus de noi.

În final, ar fi interesantă analizarea și compararea abilității de economisire a metodelor de detectare pentru care autorii nu au comunicat performanța în modul prezentat în acest studiu și au cuantificat performanța metodelor cu ajutorul tehnicilor ROC sau AUROC.

Limite ale demersului științific

Cel puțin trei restricții majore pot fi găsite în acest studiu. În primul rând, limitarea doar la câteva cuvinte cheie folosite în identificarea publicațiilor din baza de date Web of Science, editate între 1990 și 2020. În plus, publicațiile pe care le-am căutat erau scrise doar în limba engleză.

În al doilea rând, cercetarea a fost bazată pe ipoteza că poate fi determinată întotdeauna cu certitudine dacă o cerere este frauduloasă sau nu. Totuși, cum am prezentat anterior, câțiva cercetători au constatat că experții care prelucrează diferite dosare de daune auto au percepții ambigue despre încadrarea unei dosar ca fraudulos.

În final, și faptul că performanța metodelor de detectare a fraudei auto prezentate (sensibilitatea și specificitatea) erau testate pe seturi de date diferite este o limitare a studiului. Ar fi ideal analizarea sensibilității și specificității fiecărei metode de detectare prezentate pe aceeași set de date dezechilibrat din lumea reală și numai după aceasta trecerea la determinarea abilității de economisire.

Bibliografie

- Abdallah, A., Maarof, M. A. & Zainal, A., 2016. Fraud detection system: A survey. *Journal of Network and Computer Applications*, Volume 68, pp. 90-113.
- Achim, M. V. & Borlea, S. N., 2020. *Economic and Financial Crime: Corruption, Shadow Economy, and Money Laundering*. s.l.:Springer Nature.
- Agresti, A., 1990. *Categorical Data Analysis, Wiley Series in Probability and Mathematical Statistics*. New York: Wiley.
- Agyemang, M., Barker, K. & Alhaji, R., 2006. A comprehensive survey of numeric and symbolic outlier mining techniques. *Intelligent Data Analysis*, 10(6), pp. 521-538.
- Ahmed, S. R., 2004. *Applications of data mining in retail business*. s.l., International Conference on Information Technology: Coding and Computing, pp. 455-459.
- ALFA, 2020. *Insurance Fraud Agency*. [Online] Available at: <https://www.alfa.asso.fr/> [Accessed 19 03 2020].
- Artís, M., Ayuso, M. & Guillen, M., 1999. Modelling different types of automobile insurance fraud behaviour in the Spanish market. *Insurance: Mathematics and Economics*, 24(1-2), pp. 67-81.
- Artís, M., Ayuso, M. & Guillén, M., 2002. Detection of automobile insurance fraud with discrete choice models and misclassified claims. *Journal of Risk and Insurance*, 69(3), pp. 325-340.
- ASF norm no. 20, 2018. *Financial Supervisory Authority*. [Online] Available at: https://asfromania.ro/files/Asigurari/norme/2019/Norma%2020_2018_forma_aplicabila_dec2019.pdf [Accessed 10 10 2019].
- Association of Certified Fraud Examiners, 2019. *Insurance Fraud Handbook*, Austin: ACFE.
- Badriyah, T., Rahmaniah, L. & Syarif, I., 2018. *Nearest neighbour and statistics method based for detecting fraud in auto insurance..* s.l., IEEE, pp. 1-5.
- Bahnsen, A. C., Aouada, D. & Ottersten, B., 2014. *Bahnsen, A. C., Aouada, D., & Ottersten, B. (2014, December). Example-dependent cost-sensitive logistic regression for credit scoring..* IEEE, In Proceedings of the thirteenth international conference on machine learning and applications (ICMLA).
- Belhadji, E. B., Dionne, G. & Tarkhani, F., 2000. A model for the detection of insurance fraud. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 25(4), pp. 517-538.
- Belhadji, E.-B. & Dionne, G., 1998. Development of an expert system for the automatic detection of automobile insurance fraud. *Available at SSRN 134768*.

- Benetton, N., 1994. *A brief history of motor insurance*. [Online] Available at: <https://insurance-times.net/article/brief-history-motor-insurance> [Accessed 13 06 2020].
- Bermúdez, L. et al., 2008. A Bayesian dichotomous model with asymmetric link for fraud in insurance. *Insurance: Mathematics and Economics*, 42(2), pp. 779-786.
- Berry, M. J. & Linoff, G. S., 2004. *Data mining techniques: for marketing, sales, and customer relationship management*. Second ed. New York: Wiley.
- Bhattacharyya, S., Jha, S., Tharakunnel, K. & Westland, J. C., 2011. Data mining for credit card fraud: A comparative study. *Decision Support Systems*, 50(3), pp. 602-613.
- Bhowmik, R., 2011. Detecting auto insurance fraud by data mining techniques. *Journal of Emerging Trends in Computing and Information Sciences*, 2(4), pp. 156-162.
- Bolton, R. J. & Hand, D. J., 2002. Statistical fraud detection: A review. *Statistical science*, 17(3), pp. 235-249.
- Bornmann, L., Butz, A. & Wohlrabe, K., 2018. What are the top five journals in economics? A new meta-ranking. *Applied Economics*, 50(6), pp. 659-675.
- Brennan, P., 2012. *A comprehensive survey of methods for overcoming the class imbalance problem in fraud detection*. Dublin: Institute of technology Blanchardstown.
- Brockett, P., Derrig, R. A., Levine, L. L. & Alpert, M., 2002. Fraud classification using principal component analysis of RIDITs. *Journal of Risk and Insurance*, 69(3), pp. 341-371.
- Brockett, P. L., Xia, X. & Derrig, R. A., 1998. Using Kohonen's self-organizing feature map to uncover automobile bodily injury claims fraud. *Journal of Risk and Insurance*, 65(2), pp. 245-274.
- Caudill, S. B., Ayuso, M. & Guillén, M., 2005. Fraud detection using a multinomial logit model with missing information. *Journal of Risk and Insurance*, 72(4), pp. 539-550.
- Cecchini, M., Aytug, H., Koehler, G. J. & Pathak, P., 2010. Making words work: Using financial text as a predictor of financial events. *Decision Support Systems*, 50(1), pp. 164-175.
- Coalition Against Insurance Fraud, 2019. *The State of Insurance Fraud Technology*. [Online] Available at: https://www.insurancefraud.org/downloads/Fraud_tech_study_2019.pdf [Accessed July 2020].
- Decebal, B. M., 2005. *Controlul în asigurări*. Cluj-Napoca: Casa Cărții de Știință.
- Derrig, R. A., 2002. Insurance fraud. *Journal of Risk and Insurance*, 69(3), pp. 271-287.
- Derrig, R. A. & Ostaszewski, K. M., 1995. Fuzzy techniques of pattern recognition in risk and claim classification. *Journal of Risk and Insurance*, 62(3), pp. 447-482.

- Domingos, P., 1999. *A general method for making classifiers cost-sensitive*. s.l., Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining.
- Dong, W. et al., 2014. *The Detection of Fraudulent Financial Statements: an Integrated Language Model*. s.l., Pacific Asia Conference on Information Systems (PACIS), pp. 383-400.
- Duda, R. O., Hart, P. E. & Stock, E. G., 2001. *Pattern Classification*. New York: Wiley.
- Duman, E. & Ozelik, M. H., 2011. Detecting credit card fraud by genetic algorithm and scatter search. *Expert Systems with Applications*, 38(10), pp. 13057-13063.
- Eick, S. G. & Fyock, D. E., 1996. Visualizing corporate data. *AT&T technical journal*, 75(1), pp. 74-86.
- Ennas, G., Biggio, B. & Di Guardo, M. C., 2015. Data-driven journal meta-ranking in business and management. *Scientometrics*, 105(3), pp. 1911-1929.
- Farquad, M. A. H., Ravi, V. & Raju, S. B., 2012. Analytical CRM in banking and finance using SVM: a modified active learning-based rule extraction approach. *International Journal of Electronic Customer Relationship Management*, 6(1), pp. 48-73.
- FBI, 2011. *Financial Crimes Report 2010-2011*. [Online] Available at: <https://www.fbi.gov/stats-services/publications/financial-crimes-report-2010-2011> [Accessed 27 10 2018].
- FBI, 2012. *Insurance Fraud*. [Online] Available at: <https://www.fbi.gov/stats-services/publications/insurance-fraud> [Accessed 30 11 2019].
- Frank, B., 2018. Element of motor insurance cover. In: J. U. Ntiedo, A. A. Ikechukwu & B. U. Uduak, eds. *Anatomy of Finance*. Uyo: Department of Banking and Finance University of Uyo, pp. 342-349.
- Fraser, I. A., Hatherly, D. & Lin, K. Z., 1997. An empirical investigation of the use of analytical review by external auditors. *British Accounting Review*, Volume 29, pp. 35-48.
- Gao, Z. & Mao, Y., 2007. A framework for data mining-based anti-money laundering research. *Journal of Money Laundering Control*, 10(2), p. 170-179.
- Gepp, A., Wilson, H. J., Kumar, K. & Bhattacharya, S., 2012. A comparative analysis of decision trees vis-a-vis other computational data mining techniques in automotive insurance fraud detection. *Journal of Data Science*, 10(3), pp. 537-561.
- Ghezzi, S. G., 1983. A private network of social control: Insurance investigation units. *Social Problems*, 30(5), pp. 521-530.
- Glancy, F. & Yadav, S., 2011. A computational model for financial reporting fraud detection. *Decision Support Systems*, 50(3), pp. 595-601.

- Halvaiee, N. S. & Akbari, M. K., 2014. A novel model for credit card fraud detection using Artificial Immune Systems. *Applied soft computing*, Volume 24, pp. 40-49.
- Han, J., Pei, J. & Kamber, M., 2011. *Data mining: concepts and techniques*. Third ed. Waltham: Morgan Kaufmann Publishers.
- Hassan, A. K. I. & Abraham, A., 2016. *Modeling insurance fraud detection using imbalanced data classification*. Cham, Springer, pp. 117-127.
- He, H., Bai, Y., Garcia, E. & Li, S., 2008. *ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning*. s.l., In 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence), pp. 1322-1328.
- He, H., Wang, J., Graco, W. & Hawkins, S., 1997. Application of neural networks to detection of medical fraud. *Expert Systems with Applications*, 13(4), pp. 329-336.
- Holton, C., 2009. Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multi-billion dollar problem.. *Decision Support Systems*, 46(4), pp. 853-864.
- Hoogs, B., Kiehl, T., Lacombe, C. & Senturk, D., 2007. A genetic algorithm approach to detecting temporal patterns indicative of financial statement fraud. *Intelligent Systems in Accounting, Finance & Management: International Journal*, 15(1-2), pp. 41-56.
- Huang, S. Y., 2013. Fraud detection model by using support vector machine techniques. *International Journal of Digital Content Technology and its Applications*, 7(2), pp. 32-42.
- Humpherys, S. L. et al., 2011. Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems*, 50(3), pp. 585-594.
- IAATI, 2020. *International Association of Auto Theft Investigators*. [Online] Available at: <https://www.iaati.org/> [Accessed 10 02 2020].
- Insurance Europe, 2019. *Insurance fraud: not a victimless crime*, Brussels: Insurance Europe.
- Insurance Fraud Bureau, 2015. *Cutting corners to get cheaper motor insurance backfiring on thousands of motorists warns the ABI*. [Online] Available at: <https://www.insurancefraudbureau.org/media-centre/news/2015/cutting-corners-to-get-cheaper-motor-insurance-backfiring-on-thousands-of-motorists-warns-the-abi/> [Accessed 01 09 2018].
- IFB, 2020. *Insurance Fraud Bureau*. [Online] Available at: <https://insurancefraudbureau.org/> [Accessed 17 01 2020].
- IFED, 2020. *The Insurance Fraud Enforcement Department*. [Online] Available at: <https://www.cityoflondon.police.uk/police-forces/city-of-london-police/areas/city-of-london/about-us/about-us/ifed/> [Accessed 23 01 2020].

IPD, 2020. *Insurance & Pension Denmark*. [Online] Available at: <https://www.forsikringogpension.dk/en/> [Accessed 11 02 2020].

Insurance Sweden, 2020. *Insurance Sweden*. [Online] Available at: <https://www.svenskforsakring.se/en/> [Accessed 27 02 2020].

Jans, M., Van Der Werf, J. M., Lybaert, N. & Vanhoof, K., 2011. A business process mining application for internal transaction fraud mitigation. *Expert Systems with Applications*, 38(10), pp. 13351-13359.

Japkowicz, N., 2000. *The class imbalance problem: Significance and strategies*. s.l., Proc. of the Int'l Conf. on Artificial Intelligence.

Japkowicz, N. & Stephen, S., 2002. The class imbalance problem: A systematic study. *Intelligent data analysis*, 6(5), pp. 429-449.

Kaminski, J. L., 2004. *Insurance Fraud, OLR Research Report*. [Online] Available at: <http://www.cga.ct.gov/2005/> [Accessed 17 6 2019].

Karamizadeh, F. & Zolfagharifar, S. A., 2016. Using the clustering algorithms and rule-based of data mining to identify affecting factors in the profit and loss of third party insurance, insurance company auto. *Indian Journal of Science and Technology*, 9(7), pp. 1-9.

Kenyon, D. & Eloff, J. H., 2017. *Big data science for predicting insurance claims fraud*. Johannesburg, IEEE, pp. 40-47.

Kirkos, E., Spathis, C. & Manolopoulos, Y., 2007. Data mining techniques for the detection of fraudulent financial statements. *Expert systems with applications*, 32(4), pp. 995-1003.

Kou, Y., Lu, C.-T., Sirwongwattana, S. & Huang, Y.-P., 2004. *Survey of fraud detection techniques*. s.l., IEEE, pp. 749-754.

Kubat, M., Holte, R. C. & Matwin, S., 1998. Machine learning for the detection of oil spills in satellite radar images. *Machine learning*, 30(2-3), pp. 195-215.

Law no. 132, 2017. *Financial Supervisory Authority*. [Online] Available at: https://asfromania.ro/files/asigurari/legi/Lege%20132_2017_forma%20aplicabila_dec2019.pdf [Accessed 10 10 2018].

Law no. 136, 1995. *Portal Legislativ*. [Online] Available at: <http://legislatie.just.ro/Public/DetaliiDocument/6381> [Accessed 19 01 2020].

Lichtenthaler, U., 2018. The world's most innovative companies: a meta-ranking.. *ournal of Strategy and Management*, 11(4), pp. 497-511.

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*, Volume 70, pp. 1000-1009.

Long, B., 2005. *Automobile Insurance - A Brief History*. [Online] Available at: <https://web.archive.org/web/20060618214632/http://www.drbilllong.com/Insurance/HistoryI.html> [Accessed 07 01 2019].

Loughran, D. S., 2005. Deterring fraud: the role of general damage awards in automobile insurance settlements. *Journal of Risk and Insurance*, 72(4), pp. 551-575.

Lynn, B. D., 1974. *General insurance (Irwin series in insurance and economic security)*. 9th edition ed. Homewood(Illinois): R. D. Irwin.

Maes, S., Tuyls, K., Vanschoenwinkel, B. & Manderick, B., 2002. *Credit card fraud detection using Bayesian and neural networks*. s.l., Proceedings of the 1st international naiso congress on neuro fuzzy technologies, pp. 261-270.

Majhi, S. K., Bhattacharya, S., Pradhan, R. & Biswal, S., 2019. Fuzzy clustering using salp swarm algorithm for automobile insurance fraud detection. *Journal of Intelligent & Fuzzy Systems*, 36(3), pp. 2333-2344.

Massachusetts Regulation, 1993. *Mass.gov - Division of Insurance Regulations*. [Online] Available at: <https://www.mass.gov/service-details/division-of-insurance-regulations> [Accessed 12 12 2019].

Moldoveanu, N., 2002. *Combaterea fraudei in asigurari*. Bucuresti: Bren.

Motor Insurance Directive, 2009. *The European Parliament and the Council of the European Union*. [Online] Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32009L0103&from=en> [Accessed 15 10 2018].

Nagrecha, S., Johnson, R. A. & Chawla, V. N., 2018. FraudBuster: Reducing fraud in an auto insurance market. *Big Data*, 6(1), pp. 3-12.

National Health Care Anti-fraud Association, 2009. *Consumer Alert: The Impact of Health Care Fraud on You*. [Online] Available at: www.nhcaa.org [Accessed 17 2 2020].

Ngai, E. W. et al., 2011. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), pp. 559-569.

Nian, K. et al., 2016. Auto insurance fraud detection using unsupervised spectral ranking for anomaly. *The Journal of Finance and Data Science*, 2(1), pp. 58-75.

OGU 54, 2016. *Financial Supervisory Authority*. [Online] Available at: https://asfromania.ro/files/asigurari/legi/Oug%2054_2016.pdf [Accessed 25 05 2019].

Olszewski, D., 2014. Olszewski, Dominik. "Fraud detection using self-organizing map visualizing the user profiles. *Knowledge-Based Systems* , Volume 70, pp. 324-334.

Oxford Languages, 1999. *Oxford Concise English Dictionary*. Tenth ed. Oxford: Oxford University Press.

Padmaja, T. M., Dhulipalla, N., Bapi, R. S. & Krishna, P. R., 2007. *Unbalanced data classification using extreme outlier elimination and sampling techniques for fraud detection*. s.l., IEEE, pp. 511-516.

Panigrahi, S., Kundu, A., Sural, S. & Majumdar, A. K., 2009. Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning. *Information Fusion*, 10(4), pp. 354-363.

Pathak, J., Vidyarthi, N. & Summers, S. L., 2005. A fuzzy-based algorithm for auditors to detect elements of fraud in settled insurance claims. *Managerial Auditing Journal*, 20(6), pp. 632-644.

Peng, Y. et al., 2006. *Application of clustering methods to health insurance fraud detection*. s.l., 2006 International Conference on Service Systems and Service Management, pp. 116-120.

Pérez, J. M. et al., 2005. *Consolidated tree classifier learning in a car insurance fraud detection domain with class imbalance*. Berlin, Springer, pp. 381-389.

Phua, C., Alahakoon, D. & Lee, V., 2004. Minority report in fraud detection: classification of skewed data. *Acm sigkdd explorations newsletter*, 6(1), pp. 50-59.

Pinquet, J., Ayuso, M. & Guillén, M., 2007. Selection bias and auditing policies for insurance claims. *Journal of Risk and Insurance*, 74(2), pp. 425-440.

Powers, D. M., 2010. *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation*. s.l.:arXiv preprint arXiv:2010.16061.

Powers, D. M., 2012. *The problem with kappa*. s.l., Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics.

Priyanka, D. & Sri, H. B., 2019. *Analysis of Various Fraud Behaviors Using Soft Computing Techniques*. Singapore, Springer, pp. 601-614.

Quah, J. T. & Sriganesh, M., 2008. Quah, J. T., & Sriganesh, M. (2008). Real-time credit card fraud detection using computational intelligence. *Expert systems with applications*, 35(4), pp. 1721-1732.

Ravisankar, P., Ravi, V., Rao, R. G. & Bose, I., 2011. Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems*, 50(2), pp. 491-500.

Remi, O., 2006. *Elements of Insurance: Ofas Investment Limited*. Yaba: Lagos.

Road Traffic Act, 1930. *Legislation.gov.uk*. [Online] Available at: https://www.legislation.gov.uk/ukpga/1930/43/pdfs/ukpga_19300043_en.pdf [Accessed 05 07 2020].

- Safta, I. L., Achim, M. V. & Borlea, S. N., 2020. Manipulation of Financial Statements Through the Use of Creative Accounting. Case of Romanian Companies. *Studia Universitatis „Vasile Goldis” Arad–Economics Series*, 30(3), pp. 90-107.
- Sahin, Y., Bulkan, S. & Duman, E., 2013. A cost-sensitive decision tree approach for fraud detection. *Expert Systems with Applications*, 40(15), pp. 5916-5923.
- Sánchez, D., Vila, M. A., Cerda, L. & Serrano, J. M., 2009. Association rules applied to credit card fraud detection. *Expert systems with applications*, 36(2), pp. 3630-3640.
- Sharma, S., 1996. *Applied Multivariate Techniques*. New York: Wiley.
- Shaw, M. J., Subramaniam, C., Tan, G. W. & Welge, M. E., 2001. Knowledge management and data mining for marketing. *Decision support systems*, 31(1), pp. 127-137.
- Sims, M. B., 2019. *The Different Types of Auto Insurance Coverage Explained*. [Online] Available at: <https://www.autoinsurance.org/types-of-auto-insurance-coverage/> [Accessed: 27 10 2019].
- Sohl, J. E. & Venkatachalam, A. R., 1995. A neural network approach to forecasting model selection. *Information & Management*, 29(6), pp. 297-303.
- Sternberg, M. & Reynolds, R., 1997. Using cultural algorithms to support re-engineering of rule-based expert systems in dynamic performance environments: a case study in fraud detection.. *IEEE Transactions on Evolutionary Computation*, 1(4), pp. 225-243.
- Šubelj, L., Furlan, Š. & Bajec, M., 2011. An expert system for detecting automobile insurance fraud using social network analysis. *Expert Systems with Applications*, 38(1), pp. 1039-1052.
- Subudhi, S. & Panigrahi, S., 2017. 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), pp. 568-575.
- Subudhi, S. & Panigrahi, S., 2018. *Effect of Class Imbalanceness in Detecting Automobile Insurance Fraud*. s.l., IEEE, pp. 528-531.
- Sundarkumar, G. G. & Ravi, V., 2015. A novel hybrid undersampling method for mining unbalanced datasets in banking and insurance. *Engineering Applications of Artificial Intelligence*, Volume 37, pp. 368-377.
- Sundarkumar, G. G., Ravi, V. & Siddeshwar, V., 2015. *One-class support vector machine based undersampling: Application to churn prediction and insurance fraud detection*. s.l., IEEE, pp. 231-237.
- Talanx Gruppe, 2014. *Germany's law on compulsory motor insurance marks its 75th anniversary*. [Online] Available at: https://www.talanx.com/newsroom/aktuelle-themen/2014/11_geb_kfz_pflchtversicherung.aspx?sc_lang=en [Accessed 12 03 2019].

- Tan, P., Steinbach, M. & Kumar, V., 2005. *Introduction to data mining*. First ed. s.l.:Addison-Wesley Longman Publishing.
- Tao, H., Zhixin, L. & Xiaodong, S., 2012. *Insurance fraud identification research based on fuzzy support vector machine with dual membership*. s.l., IEEE, pp. 457-460.
- Tennyson, S. & Salsas-Forn, P., 2002. Claims auditing in automobile insurance: fraud detection and deterrence objectives. *Journal of Risk and Insurance*, 69(3), pp. 289-308.
- Thomson Reuters, 2009. *Usa healthcare fraud report*, New York: Thomson Reuters.
- Turban, E., Aronson, J. E., Liang, T. P. & Sharda, R., 2007. *Decision Support and Business Intelligence Systems*. Eighth ed. s.l.:Pearson Education.
- UNESPA, 2019. *Spanish Union of Insurance and Reinsurance Companies*. [Online] Available at: <https://www.unespa.es/en/> [Accessed 15 11 2019].
- Vasu, M. & Ravi, V., 2011. A hybrid under-sampling approach for mining unbalanced datasets: applications to banking and insurance. *International Journal of Data Mining, Modelling and Management*, 3(1), pp. 75-105.
- Vatsa, V., Sural, S. & Majumdar, A. K., 2005. *Vatsa, Vishal, Shamik Sural, and Arun K. Majumdar. "A game-theoretic approach to credit card fraud detection*. Berlin, Springer, pp. 263-276.
- Vaughan, E. J. & Vaughan, T., 2007. *Fundamentals of risk and insurance*. New York: John Wiley & Sons.
- Viaene, S. et al., 2007. Strategies for detecting fraudulent claims in the automobile insurance industry. *European Journal of Operational Research*, 176(1), pp. 565-583.
- Viaene, S., Dedene, G. & Derrig, R. A., 2005. Auto claim fraud detection using Bayesian learning neural networks. *Expert Systems with Applications*, 29(3), pp. 653-666.
- Viaene, S., Derrig, D. A. & Dedene, G., 2004. A case study of applying boosting Naive Bayes to claim fraud diagnosis. *EEE Transactions on Knowledge and Data Engineering*, 16(5), pp. 612-620.
- Viaene, S., Derrig, R. A., Baesens, B. & Dedene, G., 2002. A comparison of state-of-the-art classification techniques for expert automobile insurance claim fraud detection. *Journal of Risk and Insurance*, 69(3), pp. 373-421.
- Wang, J.-H., Liao, Y. L., Tsai, T. M. & Hung, G., 2006. *Technology-based financial frauds in Taiwan: issues and approaches*. s.l., IEEE International Conference on Systems, Man and Cybernetics, pp. 1120-112.
- Wang, Y. & Xu, W., 2018. Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decision Support Systems*, Volume 105, pp. 87-95.

- Warren, D. E. & Schweitzer, M. E., 2018. When lying does not pay: How experts detect insurance fraud. *Journal of Business Ethics*, 150(3), pp. 711-726.
- Webb, A., 2003. *Statistical Pattern Recognition*. Second ed. Chichester: John Wiley & Sons.
- Weisberg, H. I. & Derrig, R. A., 1991. Fraud and Automobile Insurance: A Report on Bodily Injury Liability Claims in Massachusetts. *Journal of Insurance Regulation*, 9(4), pp. 497-541.
- Weisberg, H. I. & Derrig, R. A., 1998. Quantitative methods for detecting fraudulent automobile bodily injury claims. *Risques*, 35(July–September), pp. 75-99.
- Wenzel, T., 1995. *Analysis of national pay-as-you-drive insurance systems and other variable driving charge*. [Online] Available at: <https://www.osti.gov/biblio/125357> [Accessed 10 12 2018].
- West, J. & Bhattacharya, M., 2016. Intelligent financial fraud detection: a comprehensive review. *Computers & Security*, Volume 57, pp. 47-66.
- Whitrow, C. et al., 2009. Transaction aggregation as a strategy for credit card fraud detection. *Data Mining and Knowledge Discovery*, 18(1), pp. 30-55.
- Wilson, J. H., 2009. An analytical approach to detecting insurance fraud using logistic regression. *Journal of Finance and Accountancy*, 85(150), pp. 1-15.
- Wong, N., Ray, P., Stephens, G. & Lewis, L., 2012. Artificial immune systems for the detection of credit card fraud: an architecture, prototype and preliminary results. *Information Systems Journal*, 22(1), pp. 53-76.
- Wu, S. X. & Banzhaf, W., 2008. *Combatting financial fraud: a coevolutionary anomaly detection approach*. s.l., ACM, pp. 1673-1680.
- Xu, L. & Chow, M. Y., 2006. A classification approach for power distribution systems fault cause identification.. *IEEE Transactions on Power Systems*, 21(1), pp. 53-60.
- Xu, W., Wang, S., Zhang, D. & Yang, B., 2011. *Random rough subspace based neural network ensemble for insurance fraud detection*. s.l., IEEE, pp. 1276-1280.
- Yamanishi, K., Takeuchi, J. I., Williams, G. & Milne, P., 2004. On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms. *Data Mining and Knowledge Discovery*, 8(3), pp. 275-300.
- Yan, C. & Li, Y., 2015. *The Identification Algorithm and Model Construction of Automobile Insurance Fraud Based on Data Mining*. s.l., IEEE, pp. 1922-1928.
- Yue, D. et al., 2007. *A review of data mining-based financial fraud detection research*. s.l., IEEE, pp. 5519-5522.
- Zaki, M. & Theodoulidis, B., 2013. Analyzing financial fraud cases using a linguistics-based text mining approach.

Zelenkov, Y., 2019. Example-dependent cost-sensitive adaptive boosting. *Expert Systems with Applications*, Volume 135, pp. 71-82.

Zhang, D. & Zhou, L., 2004. Discovering golden nuggets: data mining in financial application. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 34(4), pp. 513-522.

Zhang, G., Patuwo, B. E. & Hu, M. J., 1998. Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, 14(1), pp. 35-62.

Zhou, W. & Kapoor, G., 2011. Detecting evolutionary financial statement fraud. *Decision Support Systems*, 50(3), pp. 570-575.

Zhou, Z. H. & Liu, X. Y., 2005. Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Transactions on knowledge and data engineering*, 18(1), pp. 63-77.

Zimmermann, C., 2013. Academic rankings with RePEc. *Econometrics*, 1(3), pp. 249-280.

Referințe electronice

www.abcasigurari.ro
www.alfa.asso.fr
www.allianztiriatic.ro
www.asfromania.ro/
www.asirom.ro
www.baar.ro
www.cga.ct.gov
www.cityinsurance.ro
www.cityoflondon.police.uk
www.data.gov.ro
www.ergo.ro
www.euroins.ro
www.fbi.gov
www.forsikringogpension.dk/en/
www.generali.ro
www.gothaer.ro
www.groupama.ro
www.iaati.org/
www.iii.org
www.insurance-times.net
www.insuranceeurope.eu
www.insurancefraud.org
www.insurancefraudbureau.org
www.legislatie.just.ro
www.legislation.gov.uk
www.mass.gov
www.nhcaa.org
www.omniasig.ro
www.osti.gov
www.svenskforsakring.se/en/
www.talanx.com
www.unespa.es/en/
www.uniqa.ro
www.worldbank.org