



데이터 분석을 이용한 보험사기 방지

변혜원 연구위원, 손지영 연구원

- 경기침체로 인한 가계 재정 악화는 보험사기 유인을 높일 수 있으므로, 코로나19로 인한 경기침체가 지속 될 경우 보험사기가 증가할 수 있다는 우려가 제기됨
 - 최근 들어 전문적인 보험사기자가 일반 보험계약자를 보험사기로 유인하는 사례가 증가하였는데, 몇 년간 정체 또는 감소했던 국내 보험사기 적발인원이 2019년 크게 늘어남
- 최근 미국 설문조사에 의하면, 정보통신기술이 발달함에 따라 보험사기 방지방법도 미리 정해진 규칙에 따라 조사자가 의심 건을 추출하는 방법에서 데이터 분석기술을 활용하는 방향으로 이동하고 있음
 - 보험회사는 데이터 분석을 바탕으로 보험사기를 선제적으로 예방하거나 적발하고 있는데, 보험금 지급 단계뿐만 아니라 계약심사단계, 계약갱신단계에서 보험사기 방지 기술을 사용하고 있음
 - 가용데이터가 늘어나면서 보험사기 방지에 인공지능(AI)을 활용하는 사례도 증가하고 있는데, AI기술은 과거 기술에 비해 보험사기가 아닌 건을 보험사기 건으로 분류할 가능성을 감소시킨다고 평가받음
- 국내 보험회사에서는 주로 자동 위험징후, 예측모형, 데이터 시각화, 이상감지 등의 기술을 보험사기 방지·적발에 사용 중이며, 회사 내 보험금청구 및 계약 정보와 한국신용정보원 정보를 활용함
 - 한국신용정보원은 보험계약·청구·지급 데이터베이스인 보험신용정보통합조회시스템(ICIS)을 분석하여 보험사기 유의지표와 의료기관별 통계를 제공함
 - 한편 금융감독원 보험사기대응단은 보험사기인지시스템(IFAS: Insurance Fraud Analysis System)을 통해 보험계약 및 사고 등을 데이터베이스로 관리·분석함으로써 보험사기 혐의자를 추출해냄
- 이러한 노력에도 불구하고 빠르게 진화하는 보험사기를 따라잡기 위해서는 보험사기 방지·적발 기법 또한 고도화되어야 할 것임
 - 먼저 효과적인 분석을 위해서는 보험사기 관련 정보를 공유할 필요가 있으며, 이를 위해서는 공유하는 데이터의 정의나 형식을 표준화해야 할 것임
 - 또한 새로운 보험사기 데이터 분석기법이나 AI기법 활용을 적극적으로 고려하되, 데이터 집적단계에서의 데이터 품질 관리와 비정형 데이터 등 분석정보의 범위도 확대할 필요가 있음
 - 아울러 AI시스템이 부당한 차별이나 사생활침해 등의 문제를 야기하지 않도록 점검해야 할 것임

1. 검토 배경



- 경기침체로 인한 가계재정 악화는 보험사기 유인을 높일 수 있으므로, 코로나19로 인한 경기침체가 지속될 경우 보험사기가 증가할 수 있다는 우려가 제기됨
 - 영국보험협회(ABI: Association of British Insurers)는 2008년 글로벌 금융위기 당시 영국 내 허위 보험청구 건수가 전년 동기 대비 17% 증가하였다고 보고함¹⁾
 - 미국 플로리다 주보험감독당국에 따르면 2008년 플로리다주 보험사기 건수가 전년 동기 대비 21% 증가한 것으로 나타나며, 보험 종목별로는 자동차 보험, 주택담보대출 보험, 산재보험 등에 집중됨²⁾
 - 세계은행(World Bank)은 코로나19의 영향으로 2020년의 세계 경제성장률이 2008년 금융위기 이후 가장 낮은 수준인 -4.9%가 될 것으로 전망하였음³⁾

- 국내에서도 전문적인 보험사기자가 일반 보험계약자를 보험사기에 공모시키는 유형의 보험사기가 증가하고 있는데, 경기침체는 이러한 현상을 더욱 악화시킬 것으로 예상됨
 - 최근 들어 SNS나 블로그에 구인광고를 가장하여 보험사기 공범을 모집하거나, 보험사기를 조장하는 글들이 게시되어 일반 보험계약자를 보험사기로 유인하는 사례가 증가하였음⁴⁾

- 실제로 국내 보험사기 적발금액은 그동안 꾸준히 증가한 반면, 몇 년간 정체 또는 감소했던 적발인원이 2019년 크게 증가함⁵⁾
 - 금융감독원에 따르면 2019년 기준으로 1인당 적발금액이 천만 원 미만인 경우가 보험사기의 약 84%를 차지하였으며, 상해·질병, 자동차사고 피해를 과장하거나 왜곡하여 보험금을 청구하는 보험사기가 증가하였음

1) BAE Systems(2020. 3. 25), "Fraud Thrives in a Crisis - Why the Insurance Community Needs to Stay Vigilant"; <https://www.baesystems.com/en/cybersecurity/blog/fraud-thrives-in-a-crisis>

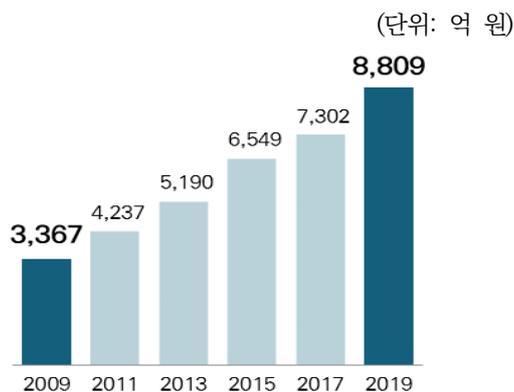
2) Florida Department of Financial Services, Division of Insurance Fraud, Crime Intelligence Analyst Unit, (2009. 8. 7), "Impact of the Economy and Insurance Fraud"

3) World Bank(2020. 6), "World Economic Outlook"

4) 금융감독원 보도자료(2020. 5. 19), "코로나19 상황을 틈타 급전 필요한 분, 고액 일당 지급 등을 미끼로 한 보험사기 주의!"

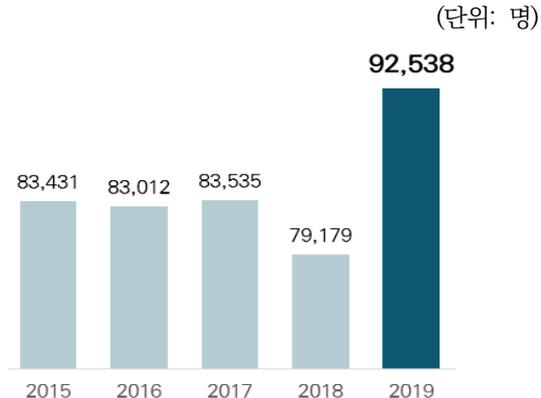
5) 금융감독원 보도자료(2020. 4. 9), "2019년 보험사기 8,809억원 적발, 전년대비 10.4% 증가"

〈그림 1〉 국내 보험사기 적발금액 추이



자료: 금융감독원

〈그림 2〉 국내 보험사기 적발인원 추이



자료: 금융감독원

■ 보험사기는 선량한 보험계약자의 보험료를 인상시킬 뿐만 아니라 사회에 사기행위를 전염시키고 사회구성원 간 신뢰를 약화시키므로, 보험사기 방지는 건전한 보험시장과 사회구성원 간의 신뢰 유지를 위해 중요함

● 예를 들어 사고피해자가 고의사고 또는 피해사실을 과장하는 경우 운전자의 보험료가 할증될 수 있는데, 2019년 중 확인된 자동차보험사기 피해자는 5,466명이었으며 약 14억 원의 보험료가 환급되었음⁶⁾

■ 본고는 국내 보험사기 대응 현황과 개선방안을 데이터 분석에 초점을 맞추어 살펴보고자 함

● 보험사기 방지를 위해서는 보험상품 개발단계 중 잠재적 보험사기 발생 가능성 최소화, 계약심사·판매·보험금 지급단계에서의 보험사기 가능성 점검 및 방지, 조기적발 및 조사를 위한 정보공유 및 데이터 분석, 보험사기 수사 효율화, 보험사기에 대한 인식 제고 등이 필요함

2. 정보통신기술 발달과 보험사기 방지



■ 보험산업은 과거부터 데이터 분석에 근거하여 사업을 영위해 왔으나, 최근 가용한 데이터가 늘어나면서 빅데이터 분석을 다양한 업무분야에 활용하는 사례가 증가함

● 빅데이터란 다양한 변수들로 구성된 데이터 세트로서 적절한 시간 내에 데이터로부터 가치를 추출해 내

6) 금융감독원 보도자료(2019. 11. 8), “보험사기 피해로 할증된 자동차보험료를 환급 받으세요”

기 위해 대규모의 연산력과 비표준적인 소프트웨어, 방법론을 사용해야 하는 데이터를 의미함
 - 일반적으로 빅데이터의 특징을 3V(Volume, Variety, Velocity)로 표현하는데, 빅데이터는 데이터
 량이 많고, 데이터 구조가 다양하고, 데이터 생성속도와 분석속도가 빠르다는 특징을 가짐

■ 보험회사는 제보를 통해 보험사기 조사·적발을 시작하지만, 관련 데이터 분석을 바탕으로 보험사기 유형이나
 보험사기자의 특성 등을 파악하여 보험사기를 선제적으로 예방하거나 적발할 수 있음

● 기존에는 보험금 지급단계나 보험사기 발생 후 조사 단계에서의 분석이 주를 이루었다면, 최근에는 보험
 계약심사단계, 계약갱신단계에서 데이터 분석을 이용하여 보험사기를 미리 예방하는 사례가 증가함

■ 미국 CAIF의 설문조사⁷⁾에 의하면 보험사기 방지·적발방법도 조사자가 미리 정해진 규칙이나 의심징후를 사용하여
 보험사기 의심 건을 추출하는 전통적인 방법에서 데이터 분석기술을 활용하는 방향으로 이동하고 있음

● 2016년에 비해 자동화된 의심징후 적발의 활용비중은 감소한 반면, 예측모형, 이상요소 감지·예외 보고,
 위치정보 분석 등의 활용비중이 증가함

- 다만, 여전히 보험사기 방지 시스템 활용기술 중 자동화된 의심징후 적발, 사회관계망분석(SNA)⁸⁾의
 비중이 가장 컸음

● 보험회사들은 보험사기 방지 시스템의 주요 장점으로서 시스템이 추출한 의심 건의 품질, 손실 경감, 조
 사 효율성 개선을 꼽음

■ 한편 사용할 수 있는 데이터양이 늘어나면서 보험사기 방지시스템에 인공지능(AI: Artificial Intelligence) 활용도
 증가하고 있는데, AI기술은 과거 적발기술에 비해 보험사기가 아닌 건을 보험사기 의심 건으로 분류할 가능성⁹⁾을
 유의하게 감소시킨다고 평가받음

● AI는 학습이나 문제해결과 같은 인간의 인지기능을 모방하는 기계 또는 컴퓨터를 이르는데, 이용할 수
 있는 데이터가 많아질수록 AI는 더 효과적으로 학습하고 판단할 수 있게 됨

● 보험사기 방지에 활용하는 AI기술에는 이상감지(Anomaly Detection), SNA, 자연언어처리,¹⁰⁾ 머신러
 ning, 음성인식, 이미지분석, 웹크롤링(Web Crawling) 등이 있음

- 머신러닝 중 비지도학습(Unsupervised Learning)은 조사자가 어떠한 특정 데이터가 보험사기를

7) Coalition Against Insurance Fraud(2019), *The State of Insurance Fraud Technology*, CAIF; CAIF(Coalition Against Insurance Fraud)는 보험사기를 방지하기 위해 1993년에 설립된 미국 소비자단체, 정부, 보험자단체 등의 연합 단체임. CAIF는 2014년부터 2년에 한 번씩 보험회사들을 대상으로 보험사기적발 기술 현황에 대한 온라인 설문조사를 실시하였는데, 2018년 조사에서는 84개 보험회사를 대상으로 설문조사를 실시함

8) Social Network Analysis

9) Type I Error 또는 False Positive라고도 함

10) Natural Language Processing(NLP)

- 보여줄 수 있는지 알지 못할 경우 보험사기 의심 건을 적발하는 데에 매우 효과적임¹¹⁾
- 음성인식 기술은 감정분석(Sentiment Analysis) 알고리즘을 사용하여 사람이 인지하기 어려운 잠재적 사기행태 지표를 식별할 수 있음

- 미국 손해보험회사를 대상으로 한 설문조사¹²⁾에 따르면 응답자의 약 56%가 보험사기 적발에 AI를 사용하고 있다고 응답하였으며, AI 활용에 있어서의 주요 고민거리로 데이터 품질과 데이터 부족을 들었음
- 응답자의 약 48%가 향후 1년간 AI에 대한 투자를 확대할 것으로 판단한다고 응답함

3. 국내 현황: 데이터 분석을 활용한 보험사기 방지



- 국내 보험회사에서는 주로 자동 위험징후, 예측모형, 데이터 시각화, 이상감지 등의 기술을 보험사기 방지·적발에 사용 중이며, 회사 내 보험금청구 및 계약 정보와 한국신용정보원 정보를 활용함
 - 회사에 따라서 데이터를 분석하여 보험사기 방지 업무를 돕는 시스템(보험사기 방지 시스템)을 회사 자체적으로 구축·운영하기도 하고, 보험사기 방지 시스템 전문회사가 구축한 시스템을 회사 내부에서 운영하기도 함
 - 보험사기 방지 시스템은 주로 자동차보험사기, 조직적 보험사기, 의료비 관련 보험사기, 언더라이팅 사기, 설계사가 연루된 보험사기 적발에 사용됨
- 보험사기 적발에 있어서 정보의 공유와 분석이 매우 중요한데, 우리나라에서 보험사기 관련 데이터를 집적하는 대표적인 기관은 한국신용정보원과 금융감독원이 있음
 - 동일한 보험사기자는 여러 보험회사를 상대로 유사한 사기를 저지르는 경우가 많으므로, 한 보험회사의 데이터만을 분석할 경우 적발이 불가능하더라도, 여러 보험회사의 데이터를 분석할 경우 동일한 수법, 동일한 사기자의 사기 패턴을 발견해 낼 수 있음

11) 머신러닝은 크게 지도학습(Supervised Learning)과 비지도학습(Unsupervised Learning)으로 분류할 수 있는데, 지도 학습이 명시적인 정답이 주어진 상태에서 컴퓨터를 학습시키는 방법이라면, 비지도학습은 명시적인 정답이 주어지지 않은 상태에서 컴퓨터를 학습시키는 방법론임

12) Coalition Against Insurance Fraud(2020), *Artificial Intelligence and Insurance Fraud*, CAIF; 동 설문조사는 미국 30개 손해보험회사를 대상으로 하여 온라인으로 조사됨

- 한국신용정보원은 보험계약·청구·지급 데이터베이스인 보험신용정보통합조회시스템(ICIS)을 분석하여 보험사기 유의지표와 의료기관별 통계를 제공함
 - ICIS는 생명보험회사, 손해보험회사, 공제기관 등 43개 회사의 보험계약 및 보험금 지급 정보를 집적하고, 동 회사들의 계약인수·보험금지급심사를 지원함
 - 보험개발원이 제공하던 사고·지급 데이터베이스인 보험사고정보시스템(ICPS)¹³⁾이 2018년말에 한국신용정보원에 이관되면서, 보험사고·보험금지급 정보도 ICIS에 통합됨
- 한편 금융감독원 보험사기대응단은 보험사기인지시스템(IFAS: Insurance Fraud Analysis System)을 통해 보험계약 및 사고 등을 데이터베이스로 관리·분석함으로써 보험사기 혐의자를 추출해냄¹⁴⁾
 - IFAS는 보험금 관련 관계의 집중도, 매개도 등을 계산하여 공모의심그룹을 추출하는 사회관계망 분석, 혐의자의 보험사기 가능성을 계량적으로 측정하는 사기지표(FI: Fraud Indicator) 등을 사용함
 - 개인단위로 산출되는 FI 76개 지표에는 보험료, 보험계약 회사 수부터 사고근접계약(보험가입 직후 보험사고 발생) 수, 사고 당 동승자 수(자동차의 경우) 등이 포함됨
 - 최근에는 금융감독원에 보고된 보험사기자들이 공통적으로 이용한 병원, 정비업체, 설계사 등을 분석, 추출하는 방법도 도입함

4. 개선방안



- 이러한 노력에도 불구하고 빠르게 진화하는 보험사기를 따라잡기 위해서는 보험사기 방지·적발 기법 또한 고도화하여야 할 것임
 - 일부 보험회사들은 AI를 접목한 보험사기 예측 시스템을 가동하기 시작했으며, 한국신용정보원도 집적된 머신러닝을 이용하여 보험사기 예측모형과 보험사기 유의지수를 개발 중임¹⁵⁾
 - 효과적인 데이터 분석과 머신러닝을 통한 적발시스템 정교화는 보험회사가 많은 양의 정보를 신속하게 분석함으로써 보험사기로 보험금이 지급되는 것을 사전에 방지할 수 있고, 새롭게 나타나는 보험사기 유

13) ICPS(Insurance Claims Pooling System)는 보험사기 예방을 목적으로 보험사고 정보를 실시간으로 제공하고 분석을 통해 보험사기 조사를 지원하는 서비스임

14) <http://insucop.fss.or.kr/fss/insucop/task03.jsp>

15) 교보생명과 현대해상 등은 AI를 접목한 보험사기인지시스템을 자체 개발하여 운영 중인데, 동 시스템들은 과거 보험사기 적발 건을 분석하여 유사한 유형의 사건이나 보험사기 의심자를 자동으로 추출해 냄

형을 빠르게 인지하고 대응하도록 할 것임

■ 먼저 데이터 분석의 효과를 높이기 위해서는 데이터의 수량과 품질 향상이 필요하므로, 보험사기 관련 정보의 공유 활성화와 공유데이터의 정의와 형식의 표준화가 필요함

- 보험회사마다 보험사기 건 또는 보험사기 혐의 건으로 분류하는 기준이 다른데, 보험회사가 금융감독원 또는 한국신용정보원에 공유하는 통계 변수에 대한 정의 재정비 등 표준화가 필요함
- 한편 의료기관, 정비업체 등과 관련된 보험금 청구건별 정보관리를 위한 표준화도 이루어져야 할 것임
 - 현재 보험사기 관련 데이터 집적 시스템이 표준화되어 있지 않아서, 추가 정보가 필요한 경우 혐의 건과 관련된 영수증 등을 바탕으로 수기로 입력해야 상황도 종종 발생함

■ 또한 새로운 보험사기 데이터 분석기법이나 시기법 활용을 적극적으로 고려하되, 보험계약, 지급데이터 이외에도 비정형데이터(Unstructured Data) 등 분석정보의 범위도 확대할 필요가 있음

- 보험사기 방지에 있어서 AI 활용의 선도적 역할을 하고 있는 회사들은 AI가 할 수 있는 일과 할 수 없는 일을 이해하는 것이 중요하며, 관련 회사나 AI 서비스를 선택할 때 도입 용이성뿐만 아니라 위험감수 정도, 보험금 지급 경험 등 보험회사의 특성을 고려해야 한다고 조언함¹⁶⁾
- 현재는 주로 회사 내부 보험계약·보험금청구 데이터, 한국신용정보원 데이터를 사용하고 있으나, 공공데이터나 미디어, 사진, 문서, 음성데이터 등 일정한 형식으로 저장되지 않은 비정형데이터 분석을 통해 시스템의 정확도를 높일 수 있을 것임
 - 보험사기 예측, 적발에 사용할 수 있는 비정형데이터에는 SNS의 글이나 사진, 손해사정사의 메모 등이 있음

■ 아울러 데이터 분석이나 시는 보험사기 방지 전문가를 대체한다기보다는 조사업무의 효과를 높이는 수단이며, 보험 회사는 AI시스템이 부당한 차별이나 사생활침해 등의 문제를 야기하지 않도록 점검해야 할 것임

- AI시스템 사용이 야기할 수 있는 잠재적 소비자피해를 최소화하기 위해서는 최근 미국 보험감독자협의회(NAIC)가 채택한 2020년 8월 보험회사의 AI 사용 관련 지침을 참고할 만함¹⁷⁾ [kiri](#)

16) Coalition Against Insurance Fraud(2020)

17) https://content.naic.org/article/news_release_naic_unanimously_adopts_artificial_intelligence_guiding_principles.htm