

## 빅 데이터 프로세싱이 보험 상품에 미칠 수 있는 영향

The impact of big data processing on the insurance product

---

저자 (Authors)	강민구 Mingu Kang
출처 (Source)	<a href="#">한국정보과학회 학술발표논문집</a> , 2018.6, 2037-2039(3 pages)
발행처 (Publisher)	<a href="#">한국정보과학회</a> The Korean Institute of Information Scientists and Engineers
URL	<a href="http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07503583">http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07503583</a>
APA Style	강민구 (2018). 빅 데이터 프로세싱이 보험 상품에 미칠 수 있는 영향. 한국정보과학회 학술발표논문집, 2037-2039
이용정보 (Accessed)	한양대학교 166.***.140.13 2019/11/22 08:00 (KST)

---

### 저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

### Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

## 빅 데이터 프로세싱이 보험 상품에 미칠 수 있는 영향

강민구

## The impact of big data processing on the insurance product

Mingu Kang

Dept. Computer science, Sungkyunkwan University

## 요 약

보험업은 다가오는 IFRS17 적용에 따른 RBC 비율 악화, 시장 정체, 수익 구조 악화 등 여러 고민을 안고 있다. 그 고민의 해결방안 중 하나로 각광 받는 것이 현재 4차 산업의 핵심 기술 중 하나인 빅 데이터 프로세싱이다. 빅데이터 기술을 활용한 인슈어테크의 가능성은 무궁무진하여 현재 각 보험사에서도 활발한 연구가 이루어지고 있다. 그 중에서 기존에는 측정하거나 계량하여 모델화 할 수 없기 때문에 보험에 적용할 수 없었던 비정형 데이터를 실제로 적용하여 기존 보험 상품을 개량하는 실험을 해보고, 이러한 보험 상품의 개량과 새로운 risk 발견을 통한 신보험상품 개발이 보험업의 과제를 해결할 방안 중 하나가 될 수 있음을 보이고자 한다.

## 1. 서 론

보험업은 현재 여러 고민을 안고 있다. 우선 시장 성장이 정체 되어있다. 보험 가입률은 저조해지고 있고, 5%대 성장률은 매년 꾸준히 하락해서 현재 2%대의 성장률을 보이고 있다. 특히 퇴직 연금 시장 등으로 대표되는 생명보험시장은 손해보험보다 이러한 둔화가 뚜렷하여서, 성장 정체를 보험연구원에서 2016년부터 꾸준히 제기해 오고 있는 실정이다.

두 번째는 수익구조 악화이다. 다행히 자기자본이익률(ROE)이 상승하고 있고, 손해율이 낮아지고 있다고는 하지만 그와 동시에 각 보험사에 자동차 보험료 인하 경쟁이 과열화 되고 있다. 5년간 3조 정도로 집계되는 보험사기도 문제이다. 보험사기는 보험업을 그 막대한 금액으로 외적으로도, 그리고 가입자에게 공평성과 신뢰성에 대한 의문을 갖게 하는 등 내적으로도 괴롭히는 요소이다.

마지막으로는 새로운 회계기준의 적용이다. 다가오는 2020년에 적용할 회계기준인 IFRS17를 적용하더라도 보험회사는 보험업을 영위하기 위해 안정적인 재무안정성과 신뢰도를 보여주는 대표적인 지표인 보험금지급여력비율(RBC)을 유지해야한다. 그로 인해 안정적인 연착륙을 위해 쌓아야 하는 준비금은 매년 높아져 가고 있지만, 국내에서 자본을 확충할 수단은 유보, 유상증자, 후순위채 발행등으로 그렇게 많지 않다.

이러한 상황을 타파하기 위해 업계에서 주목하고 있는 것 중에 하나가 빅 데이터 기술[4]을 활용한 인슈어테크(I

nsurance+Technology)[1]이다. 4V 또는 5V로 불리며 크기(Volume), 다양성(Variety), 속도(Velocity), 정확성(Veracity), 가치(Value)의 빅데이터의 특징들은 ‘큰 수의 법칙’을 대전제로 이용하여 리스크 분배와 공유(Risk pooling & sharing)을 하는 보험업과 사고의 방식과 방향이 일치한다.

보험업은 일어나지 않은 사건 위험을 여태 일어났던 일들을 토대로 예측해야한다. 그걸 위해서는 많은 정보가 필요했고, 따라서 지금까지의 보험업은 3이원이라고 불리우는 이자율, 위험율, 사업비율 같이 안정적으로 많은 양의 정보를 얻을 수 있는 정형 데이터만을 사용해왔다. 그러나 4차 산업혁명이 오면서 이러한 제한이 없어졌다. 데이터 저장/관리 기술 발달과 프로세스, 로직 기술 발달로 인한 처리 복잡도 증가로 1초에 얻을 수 있는 데이터양과 컴퓨터의 연산속도는 기하급수적으로 증가하고 있다. 이에 보험사는 기존에 취급하던 정형 데이터들 외에도 기존에는 다루지 못하였던 수많은 비정형 데이터들을 수집하고 모델화하여 취급할 수 있게 되었다[1],[2]. 이는 기존 계약자들의 위험을 조금 더 세부적으로 분석하거나, 수많은 잠재 고객들이 원하는 바를 알아내고 분류하는 등 새로운 보험가능 리스크(insurable risk)를 만들거나 기존의 risk 폭을 넓힐 수도 있게 되었다는 소리이다. 이러한 발전은 보험업이 새로운 상품으로 새로운 시장을 개척할 수 있게 해줄 것이다.

본 논문은 배경을 설명한 이후, 빅데이터를 통해 얻어낸 비정형 데이터의 영향, 영향이 있을 만한 비정형 데이터 선정, 결론 및 향후 과제를 설명하겠다.

(출처: 도로교통공단, 미분류 날짜 제외)

## 2. 비정형 데이터 적용 보험 비교

우선, 빅데이터로 도출할 수 있는 비정형 데이터를 적용한 보험이 기존 보험 상품과 다른 상품이라고 말 할 수 있을 정도의 유의성을 가지고 있는지를 실제로 상품을 만들어 비교해보고자 한다.

본 논문에서 적용해볼 비정형데이터는 기상기후이다. 기상기후는 대표적인 빅데이터라고 볼릴 만큼 손색이 없을만큼 방대한 양의 정보를 가지고 있었으나, 지금까지는 이 자료들을 수집하고 데이터화하기 위한 인프라가 갖춰져 있지 않았었다. 그러나 기상청에서 인프라를 갖추고 타 부서와 협력하여 만든 자료를 제공[3]하기 시작한 이후로는 접근성이 높고 매년 기상청 주관 콘테스트가 열릴 정도로 여러 창의적인 아이디어가 무궁무진하게 나오는 분야가 되었다. 이 비정형 데이터로 여겨져서 기존에는 적용하지 않았던 변수인 기상기후를 기존 자동차 보험에 실제로 접목하여 미칠 수 있는 영향을 비교, 확인해보고자 한다.

우선적으로는 상품에 기상기후라는 데이터를 적용할 방식으로 사용자 기반 보험 방식과 유사한 방식을 사용해자. 사용자 기반 보험(Usage based insurance)이란 사용자의 행동이 보험료에 영향을 미치는 보험으로 사물인터넷등을 이용하여 급발진, 급제동, 과속 여부 등 기존에는 반영하지 않았던 변수들을 측정하고 지수로 변환하여 보험료 계산에 반영하는 보험이다. 마찬가지로, 우리도 기상기후가 사고에 미치는 영향을 보험료 계산에 넣는 방식으로 진행해보겠다.

즉, 우리가 비정형데이터의 영향을 파악하기 위해 만들 가상 상품은 운전자의 운전 당시 기상 상태에 점수를 매겨서 이에 따라 보험료를 할인, 할증하는 보험이다. 이러한 접근은 요일제 특약이나 위험 시간대의 운행 등을 반영하는 Pay when you drive 방식과 유사하다고 할 수 있다.

기존의 상품을 운전 중 사망 시 보험금 1을 지급하는 보험이라고 하고, 이 때 극단적으로 우리가 맑은 날씨에만 운전할 경우 보험료를 감소하는 보험 상품을 개발한다면 보험료와 보험금은 어떻게 될까?

<표1> 기상 상태간 사건 사고 비율 (2012~2016, 전국)

	맑음	흐림	비	안개	눈	합계
사고비율	85.90	4.29	8.18	0.16	0.80	99.33
사망률대비	1	1.89	1.27	5.16	1.32	

<표2> 기상상태별 교통사고에 따른 사망자수 (2012~2016, 전국)

기상상태	기준년도	2012	2013	2014	2015	2016	합계
합계	발생건수	223,656	215,354	223,552	232,035	220,917	1,115,514
	사망자수	5,392	5,092	4,762	4,621	4,292	24,159
	사망률(%)	2.41	2.36	2.13	1.99	1.94	2.17
맑음	발생건수	186,840	185,655	192,007	199,816	193,891	958,209
	사망자수	4,233	4,126	3,825	3,699	3,511	19,394
	사망률(%)	2.27	2.22	1.99	1.85	1.81	2.02
흐림	발생건수	10,850	9,860	9,682	9,581	7,858	47,831
	사망자수	407	386	364	364	308	1,829
	사망률(%)	3.75	3.91	3.76	3.80	3.92	3.82
비	발생건수	21,460	16,047	17,456	19,938	16,298	91,199
	사망자수	625	430	460	463	377	2,355
	사망률(%)	2.91	2.68	2.64	2.32	2.31	2.58
안개	발생건수	316	411	347	428	322	1,824
	사망자수	35	45	26	42	42	190
	사망률(%)	11.08	10.95	7.49	9.81	13.04	10.42
눈	발생건수	2,774	1,851	1,988	1,009	1,353	8,975
	사망자수	63	60	51	31	34	239
	사망률(%)	2.27	3.24	2.57	3.07	2.51	2.66

(출처 : 도로교통공단)

도로교통공단의 자료에 의하면 최근 자료 5년간 맑은 날에 일어난 사고는 전체의 약 85%이었다. 이때의 사망률을 1로 두어 기준으로 보았을 때 나머지 15%인 맑지 않은 날의 사망률은 맑을 때의 사망률 보다 최소 1.27배(비) ~ 5.16배(안개) 높음을 확인할 수 있다.

이를 기반으로 기존 상품의 보험금 지출량과 신상품의 보험금 지출량을 비율로 계산하면

$$\begin{aligned} \text{기존 보험금 지출량} &: 0.8590 \times 1 + 0.0429 \times 1.89 + 0.0818 \times 1.27 + 0.0016 \times 5.16 + 0.0080 \times 1.32 = 1.062783 \\ \text{새로운 보험금 지출량} &: 0.8590 \times 1 = 0.8590 \end{aligned}$$

가 된다.

논문에서 만들어본 가상 상품에서는 지출해야하는 보험금이 기존 보험금의 80.8% 밖에 되지 않는다. 이는 단순히 생각하면 수지상등의 원칙에 의거해서 신상품의 보험료가 기존의 보험료보다 최대 19.2%라는 파격적인 수치만큼 싸질 수 있음을 의미한다.

<표3> 기상 상태 별 차량 용도간 사건 사고 (2012~2016, 전국)

차량용도	기상상태	흐림	비	안개	눈	합계
비사업용	발생건수	6,168	12,616	269	1,090	111,293
	구성비율	1.8	3.6	0.1	0.3	
사업용	발생건수	3,660	7,977	570	1,015	352,277
	구성비율	3.3	7.2	0.5	0.9	

(출처: 도로교통공단, 미분류 날짜, 기타 차량 제외)

또한 빅데이터는 같은 기상상태 내에서도 차량의 종류, 운전자의 나이, 차량의 목적, 불쾌지수, 도로의 종류 등

다른 변수에 미치는 영향, 관계성을 파악하게 하여 변수들을 더 세분화하는 것을 가능하게 해준다. 위에서 언급된 차량 용도가 좋은 예시이다.

맑지 않은 날씨일 경우 사업용은 비사업용 보다 약 1.83배~5배의 사고율을 보여주었다. 물론 사업용 차량이 기본적으로 비사업용 차량보다 사고 비율이 높고, 사업 차량과 비사업 차량은 운전 시간대 등을 포함한 환경 자체가 다르기 때문에 이것이 단지 기상 상태만으로 인한 것이라고 단언할 수는 없지만, 이 정도의 높은 사고율 차이는 기상상태와 운전 차량 용도라는 비정형 변수들간의 상관관계가 있음을 보여주고 비정형 데이터들간의 영향과 기존 데이터와의 상관관계를 제시하고 있다고 생각한다.

### 3. 비정형 데이터 선정

빅데이터를 이용한 비정형 데이터의 영향력 판단에는 또한 기존의 로지스틱 모델을 이용한 회귀분석 말고도 딥러닝 모델을 이용할 수 있다. 특히, 딥러닝 모델 분석 방법으로는 민감도 분석(Sensitivity) 방식이 유효하다.

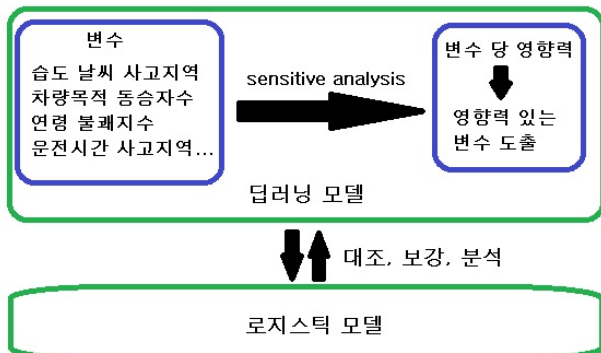


그림1. Deep learning model and sensitive analysis

민감도 분석 방식이란, 최빈값을 기준으로 잡고 영향력 있어보이는 요인 또는 변수가 변할때마다 그 변화량이 예측값에 끼치는 변동량을 측정하여 그 영향력을 재는 방식이다[6]. 그렇게 산출된 영향력은 위의 그림 1과 같이 기존의 로지스틱 모델로 산출된 변수가 영향력을 가지고 있다는 주장을 보강하거나, 변수 간 교호작용을 파악하여 그 변수를 활용한 맞춤형 상품(On-demand)을 만드는데 유용하게 사용할 수 있다.

### 4. 결론 및 향후 연구

위 실험에서 언급된 상품은 설정이 너무 극단적이고 변수간의 상관관계를 무시하는 등 단순하다. 또한, 사용자의 습관이 보험료의 할인을 결정하는 Pay how you drive 방식과는 다르게, 날씨를 이용한 Pay when you drive 방식은 본인의 의지로 수정하거나 개선이 가능한 것이 아닌 주어진 것이기 때문에 많은 가입자로 하여금 그 공정성에 의문을 가지게 하는 만큼 현실성이 떨어진다.

그러나 어디까지나 위의 실험은 보험에 계산되지 않았던 비정형 데이터가 판별하고 계산만 할 수 있다면 보험료에 유의한 영향을 줄 만한 차이를 보여줄 수 있는 새로운 변수라는 것을 보이는데 있었다. 이러한 목표는 본 논문을 통해 충분히 달성되었다고 생각한다.

빅 데이터를 실제로 적용하는 데까지 넘어야 하는 벽이 많다. 텔레매틱스 통신 장치등 신 장비 도입, 수집된 자료 모델화 및 가치 계산, 꾸준한 피드백을 통한 수정을 위한 인프라 구축, 개인정보 유출 문제, 지나친 세분화로 인한 보험사의 위험 공유(Risk pooling) 능력 약화[5], 초고위험군의 소외, 새로운 인력의 양성, 기존 대면 채널인 설계사의 실직 문제, 공공기관과 협력할 수 있는 제도 기반 마련 문제 등이 그러하다.

그러나 이번 논문 작성을 통해 이러한 벽들을 전부 넘을 만큼 빅데이터는 매력적이고 가치 있는 기술이라고 단언한다. 빅데이터가 새로운 insurable risk의 발견, 나아가 가격경쟁력을 가진 새로운 시장선도형 상품을 만들 수 있을 가능성을 충분히 제시하고 있기 때문이다. 앞으로 빅 데이터가 가져올 효익을 기대하며 논문을 마친다.

### 참고 문헌

- [1] 인슈어테크 혁명 : 현황 점검 및 과제 고찰. 2017.8 박소정. 박지윤.
- [2] KiRi 리포트, 보험산업의 빅 데이터 활용 현황 및 향후 과제. 2015-07-13, 임준,황인창,이성은
- [3] 도로교통공단 (2012~2016) TAAS 통계
- [4] 보험사, Big Data에 답을 묻다. 2014.10. 김욱.
- [5] 보험산업의 빅 데이터 활용과 사회적 위험공유, 보험연구원. 2015.08
- [6] 보험산업 Value Chain을 기준으로 개발된 빅 데이터 분석 참조모델. ING 생명보험