

World Top 10 University By 2030

AI. 보급·확산 지원 사업 제안서

AI+헬스케어+금융: 인공지능 기반 생애주기 의료비 관리

UNIST

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

산업공학과 조교수

산업공학과 조교수

연구명

AI+헬스케어+금융: 인공지능 기반 **생애주기 의료비 관리**

연구 목표

- 개인 의료 데이터를 활용한 **개인별 생애주기 의료비 예측**
- 의료비 절감을 위한 **개인 맞춤형 보험 최적화**

연구 내용

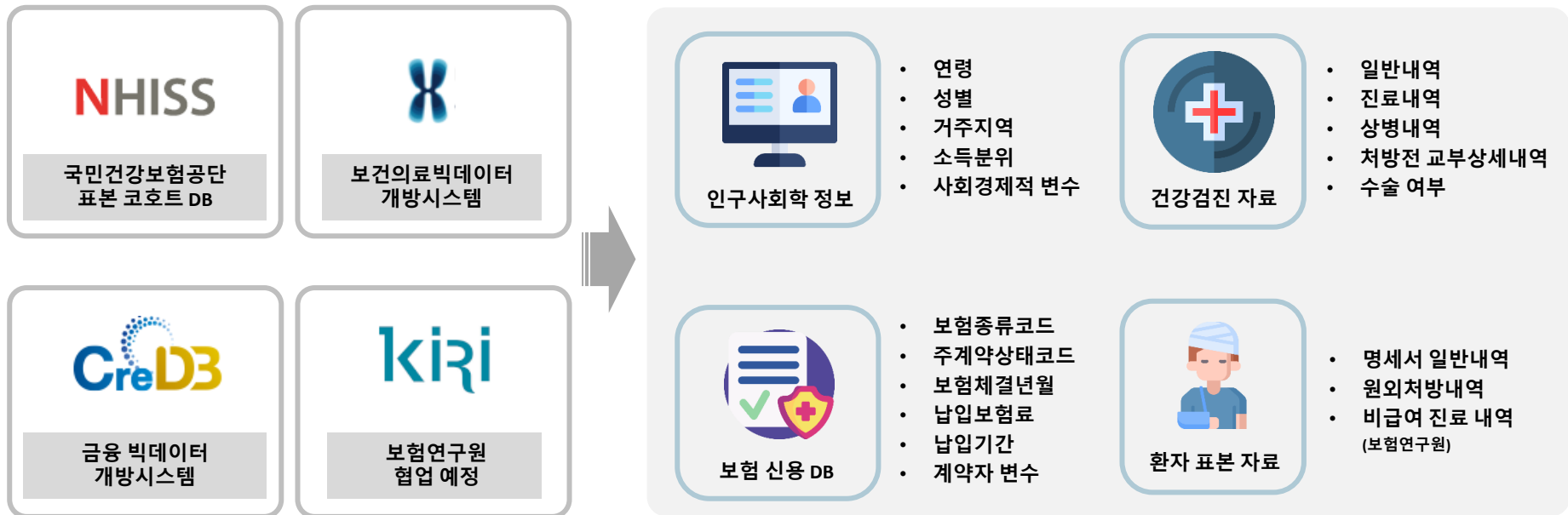
- 건강검진 데이터 기반 개인별 질병 예측 모델
- 진료 데이터 기반 환자 특성에 따른 질병별 치료비 예측 모델
- 보험 데이터 기반 보험 보장 텍스트 분석
- 개인 맞춤형 보험 최적화

기대 효과

- 개인 특성에 맞는 맞춤형 의료비 관리를 통해 국민의 건강, 금융 복지 개선
- 개인별 건강검진, 의료비, 보험 데이터를 통합적으로 고려한 첫번째 학술적 연구
- 여러 유형의 데이터를 통합하여 양질의 서비스를 제공하는 마이데이터 사업의 모범 사례

2. A.I. 적용 대상 연구

▶ 데이터 선정



1. 국민건강보험공단 표본 코호트 DB는 2017년 이후 전국민 3%인 약 145만명의 정보를 포함하고 있는 전국민을 대표하는 데이터임.
 - ✓ 환자의 1년간 의료서비스 이용 내역 포함, **명세서 일반내역, 진료내역, 상병내역, 원외처방내역 및 영양기관정보**로 구성되어 있음.
2. 보건의료빅데이터 개방시스템에는 2006년 1년간 건강보험가입자 및 의료급여수급권자 자격을 유지한 100만명을 대상의 데이터가 존재함.
 - ✓ 2002년부터 2015년까지 14년간 **의료이용현황(일반, 진료, 상병, 수술, 처방전 교부상세 등)**의 정보가 수집되어 있음.
3. 금융 빅데이터 개방 시스템은 신용정보원의 **보험신용정보 계약현황** DB에 등록된 전체 개인의 5%, 약 260만명의 데이터가 있음.
 - ✓ 계약관계자정보, 계약정보, 담보정보 데이터로 구성되어 있음.
4. 누락되어 있는 중요한 정보인 **비급여 진료 내역 데이터**는 **보험연구원**과 협업을 통해 확보할 예정임.

▶ A.I 적용 연구 소개

질병 예측 과정



성별 남성
만나이 42세
거주지역 울산광역시
직업 전문직
소득분위 3분위
건강검진자료 & 환자표본자료

질병 가능성



소화기계 질환
27.32 %



췌장 질환
17.25 %

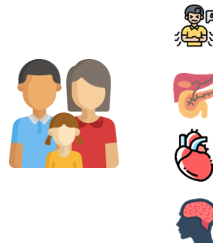


심장 혈관 질환
6.52 %



뇌 혈관 질환
1.17 %

질병별 의료비 산출



생애주기별 발생하는 개인에게
최적화된 의료비 산출

보험 추천 최적화

Advice



실손의료보험이 가능성 높은 질병에 비해 상대적으로 부족합니다.



암보험에 대해서 검토해야하는 보장내역이 존재합니다.



보장성보험이 과도하게 많이 가입되어있습니다.



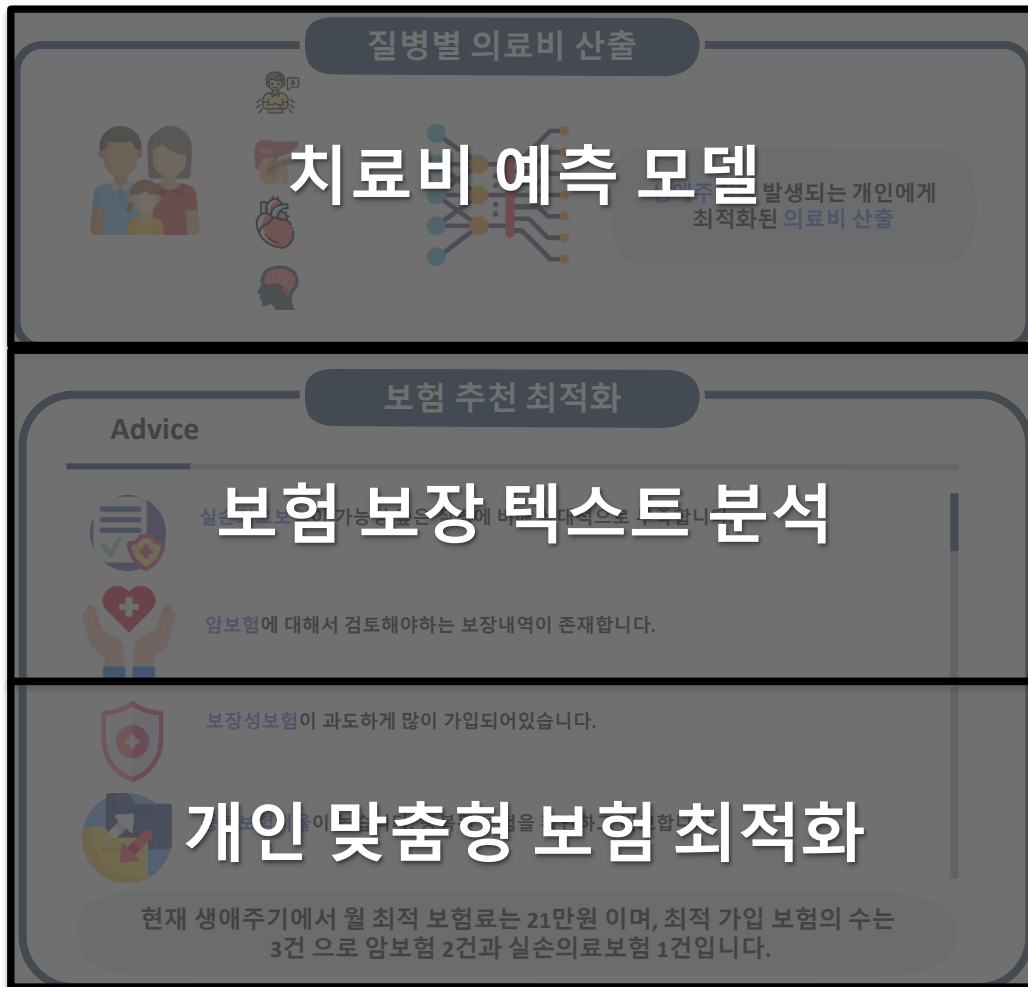
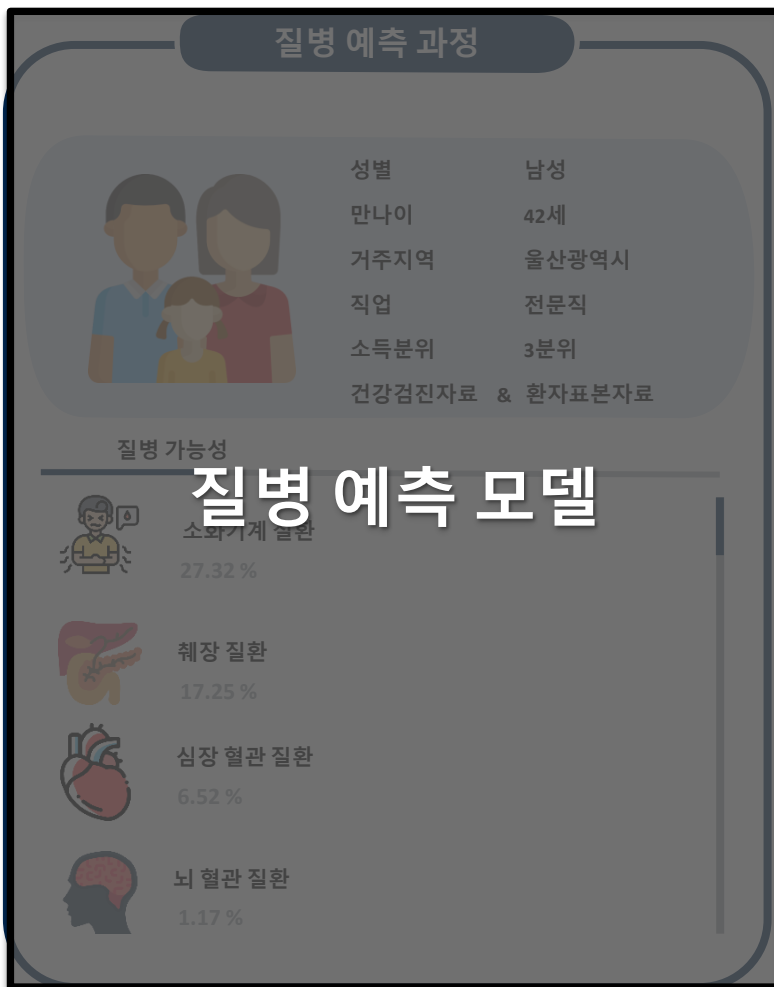
중복보험비율이 높습니다. 중복된 보험을 확인하고 비교합니다.

현재 생애주기에서 월 최적 보험료는 21만원 이며, 최적 가입 보험의 수는 3건 으로 암보험 2건과 실손의료보험 1건입니다.

- 질병 예측을 통해 각 개인에게 미래에 발생할 수도 있는 질병별 의료비를 미리 계산함으로써, **현재 상황에 최적화된 보험을 추천 가능**
- **개인의 생애주기에 따라서 발생할 수 있는 의료비 리스크를 사전에 계산하여 최적화된 보험 준비로 개인의 복리 후생을 극대화 가능**

2. A.I. 적용 대상 연구

▶ A.I 적용 연구 소개

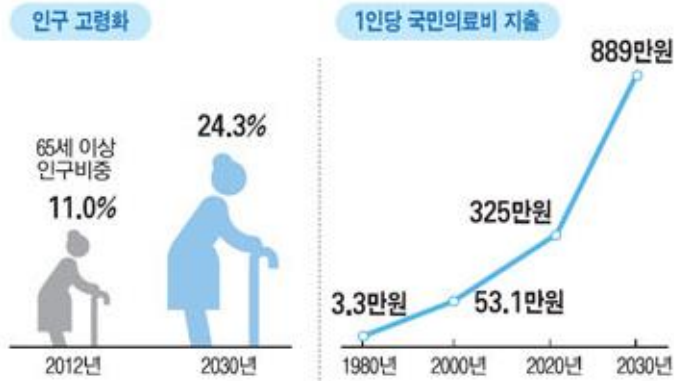


- 질병예측을 통해 각 개인에게 미래에 발생할 수도 있는 질병별 의료비를 미리 계산함으로써, 현재 상황에 최적화된 보험을 추천 가능
- 개인의 생애주기, 가족사, 건강할 수 있는 의료비 리스크를 사전에 계산하여 최적화된 보험 조합으로 개인의 부담부담을 극대화 가능

3. A.I. 적용 필요성

▶ 장수리스크와 의료비 지출의 증가

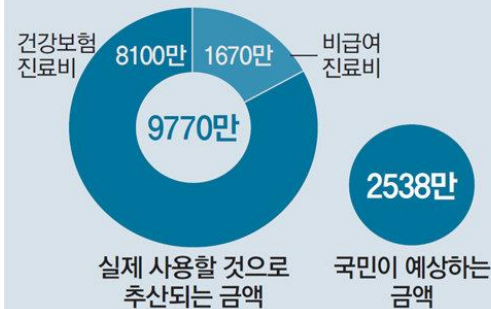
고령화와 의료비 지출



- 한국은 고령화가 세계에서 가장 빨리 진행되는 국가이며 이에 따라 의료비 지출도 급증하고 있으며 이로 인해 정부는 물론 개인의 부담도 가중될 전망

고령 의료비 설문

65세 이후 병원비 얼마 드나 단위: 원
비급여 진료비는 2014년 건강보험 대비 비율(20.6%)로 추산.



자료: 생명보험사회공헌위원회 건강보험공단

- 20대에서 50대에서 예상되는 의료비와 실제로 지출되는 의료비의 괴리가 큼.

- 가구주 연령이 높고, 만성질환(암, 뇌혈관, 심혈관 등)의 경우에 있어서 **재난적 의료비 지출은 삶의 질을 떨어뜨림.**
 - 개인정보를 바탕으로 **질병을 예측**할 수 있다면 의료비 지출의 많은 부분은 보장 받을 수 있음.
- 설문조사 결과*에 따르면 20~50대는 필요한 의료비를 실제 들어가는 금액의 3분의 1 수준으로 **턱없이 낮게 추산**하고 있음.
 - 치료비 예측이 가능하다면 예상 지출 규모를 **보다 정확히 파악할 수 있게 되며**, 이를 통해 의료비 낭비를 사전에 막을 수 있음.
- 건강은 재무적·비재무적 측면에서 노후생활에 중요한 요소로, **건강관리와 의료비 지출에 대한 대비**가 필수적임.
 - 대비가 충분하지 못하면 은퇴 후 예상치 못한 경제적 어려움에 처할 수 있음
 - **미래에 발생할 질병**에 대해 알 수 없기에 대비가 어려운 상황

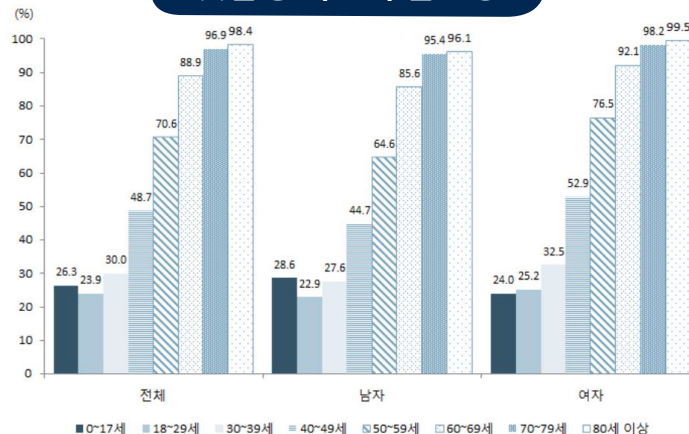
▶ 정밀의료의 중요성

의료 패러다임의 변화



- 의료에서 패러다임은 치료 중심에서 예방중심으로 이동하며, 예방 의학의 중요성이 점점 커지고 있음.
- 질병을 예방하는 것만으로도 막대한 사회적 비용을 줄일 수 있음.

맞춤형 치료의 필요성



자료 : 한국의료패널, 질환은 한국표준질병인사분류(KCD-7)에 따름.

- 연령이 증가할수록 높아지는 질병들을 개인의 내재적인 특성에 맞게 예측함으로써 사회적 비용을 줄이고 사회 후생을 증가시키므로 맞춤형 치료가 필요함.

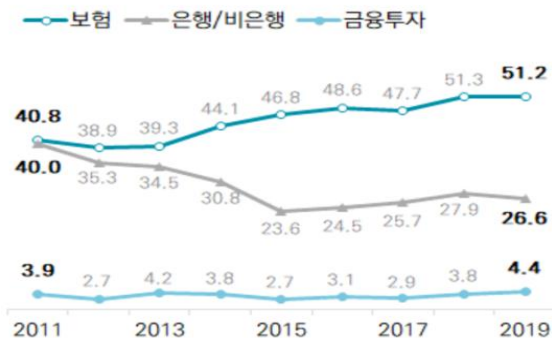
- 최근 몇 년 사이 빅데이터 시대가 열리면서 질병을 예측하고, **맞춤형 예방/치료가 가능한 세상**이 열리고 있음.
- EHR(Electronic Health Record)과 같은 진료 데이터를 분석하면 **환자별 개인 특성**에 맞는 모델을 만들 수 있음.
- 국내에서는 **정부 주도의 헬스케어 빅데이터 구축 및 활용**을 추진 중에 있음.
 - ✓ 국민건강보험공단, 건강보험심사평가원이 데이터 수집/배포 핵심 기능을 담당하고 있음.

3. A.I. 적용 필요성

▶ 보험업계의 현재 상황

금융권역별 민원건수

(단위: 천 건)



자료: 보험연구원

- 최근 8년간 보험업권의 민원건수가 크게 상승하였는데 이는 소비자의 낮은 신뢰도와 만족도를 나타냄.

보험 중복가입 현황

실손의료보험 중복가입자 추이 (단위:만명)

중복가입 유형별 현황 (단위:만명)

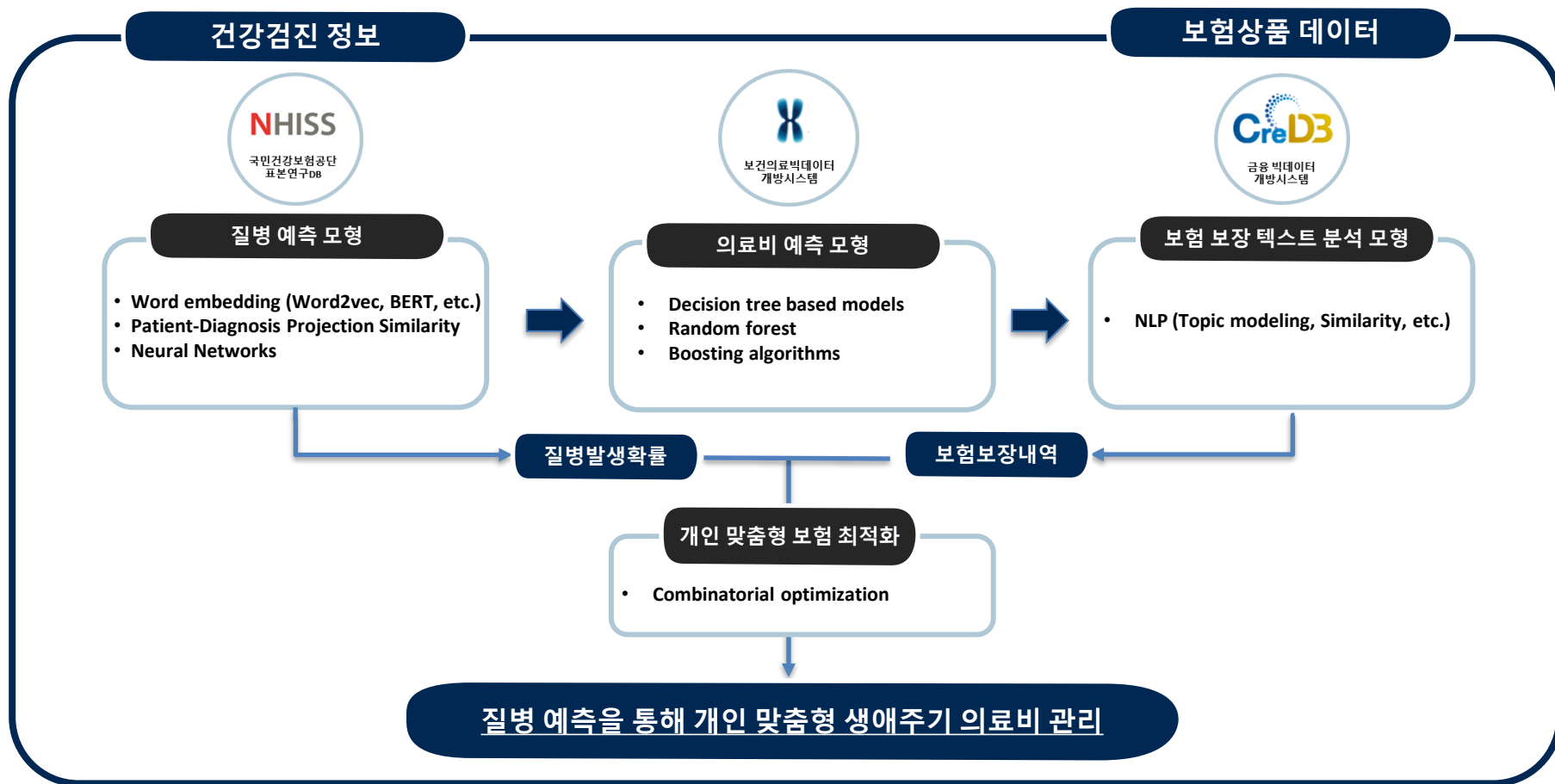


실손의료보험 중복가입자 추이, 그래픽=박구원기자

- 보험의 복잡한 보장영역, 보장기간 등으로 인해서 중복가입이 많음.
- 소비자가 잘못된 보험에 가입된 경우 이를 알려주고, 현재 상황에 맞는 보험을 추천해줘야 할 필요가 있음.

- 보험연구원에 따르면 한국에서 **보험산업의 신뢰도**는 미국, 영국 등 30개국 중 **최하위** 수준.
 - ✓ 설계사의 전문성 부족, 보험 상품의 복잡성과 과당 경쟁으로 인한 중복가입의 문제임.
- 대부분의 사람들이 보험상품에 **중복가입** 되어있는지 모른채 보험료를 납부하고 있어 **사회적 손실**이 큼.
 - ✓ 2019년 6월 기준 신용정보원에 따르면 실손보험에 중복 가입한 사람은 138만명임.
- 보험판매자와 보험가입자 간의 **역선택의 문제**가 존재하기 때문에 보험 판매자는 필요하지 않은 보험을 파는 경우가 많음.

▶ 모델 파이프라인



- 질병 및 의료비 예측 모형 단계에서는 딥러닝 기반의 자연어 처리 알고리즘과 머신러닝 기법을 활용함.
 - ✓ 80가지 이상의 다양한 질병을 한번에 예측할 수 있음.
- 보험 보장 텍스트 분석과 개인 맞춤 최적화 단계에서는 조합 최적화(Combinatorial Optimization) 기법을 활용함.
 - ✓ 개인 맞춤형 생애주기 의료비 관리를 통해 개인의 후생을 극대화 할 수 있음.

4. A.I. 기술 적용 계획 – 차별점

▶ 유사서비스 및 기존연구와의 차별점

PROBLEM

유사서비스 및 기존 연구



질병별 코호트 구축의 필요성

- 질병에 대한 예측을 위해서 질병 별로 구분된 코호트를 구축
- 예측 질환의 수가 많아질수록 코호트 구축에 많은 시간과 노력이 소요됨



시간적 영향력 반영의 한계

- 기존 방식의 코호트 구축은 특정 시기의 데이터만을 활용함
- 이는 시간적 영향력을 반영하는 것에 한계를 가져옴

SOLUTION

제안된 AI 기술



질병 구분 없이 단일 코호트 구축

- 질병 구분없이 단일 코호트 구축을 통해 질환 예측 가능
- 예측 질환의 수가 많은 경우, 기존 방식에 비해 코호트 구축 시 요구되는 시간과 노력이 절감됨



시간적 영향력 반영 가능

- 환자의 매 방문 마다의 데이터를 모두 활용하는 것이 가능함
- 이를 통해 시간적 영향력의 유의미한 반영이 가능해짐

질병 구분 없이 시간적 영향력이 반영된 단일 코호트 구축 및 이를 통한 학습이 가능

▶ 유사서비스 및 기존연구와의 차별점

PROBLEM

유사서비스 및 기존 연구



개인 맞춤형 연구의 필요성

- 의료비지출 예측 관련 연구들은 거시적인 관점에서만 살펴봄.
- 기존의 보험상품 추천은 정적인 방식으로 한계를 지님.



기존 인슈어테크 서비스

- 고객 관리와 민원프로세스를 고도화 하는 방향으로 서비스.
- 질병 예측과 동시에 보험을 추천하는 서비스는 진행되지 않음.(OSCAR, Clover, Lemonade, etc)

SOLUTION

제안된 AI 기술



개인 맞춤형 AI 모델

- 생애주기에 걸쳐 개인별로 발생 확률이 높은 질병을 예측
- 이를 바탕으로 의료비를 예측하고 맞춤형 보험 최적화



개인별 의료비 예측 및 보험 최적화

- 다양한 보험상품에 대한 개인별 보장내역 분석
- 질병 예측과 보험가입 최적화의 분리된 서비스가 아닌 결합된 서비스 제공

개인이 생애주기에서 마주할 수 있는 주요 질병을 폭 넓게 다루며 개인별 의료비 예측 및 보험 가입 최적화 가능

기대효과

- 보험산업에서 정보비대칭으로 발생하는 역선택의 문제를 해결함으로써 **보험가입자의 후생이 증가함**
- 판매채널을 이용한 각 보험사의 상품 구매 방식에서 사용자의 질병 발병 가능성과 라이프사이클에 최적화된 보험을 추천해주는 방식으로 **개인화된(Personalized) 서비스 고도화**

- 개인별 건강검진, 의료비, 보험 데이터를 통합적으로 고려한 **최초의 학술적 연구**
- 서로 다른 유형의 데이터를 통합하여 유의미한 서비스를 제공하고자 하는 **마이데이터 사업의 모범사례**

활용계획

- **뱅크샐러드, 토스와 같은 마이데이터 업체와의 협력**을 통해 모델 정교화 및 **실질적인 서비스 개발 추진**
- 근로복지공단, 건강보험공단 등 공공기관과의 협력을 통해 **공공 성격의 서비스 개발**

- 의료비 이외에도 가계금융에 중요한 요소들로 영역을 확장하여 **가계 금융을 통합적으로 진단하고 관리**할 수 있도록 할 것
- 기존의 가계금융 건강검진 모델에도 가구 구성원의 건강 관련 데이터를 추가하여 **확장할 수 있을 것**

- [1] Mongelli, Michael N., et al. "Financial burden and quality of life among thyroid cancer survivors." *Surgery* 167.3 (2020): 631-637.
- [2] Hazell, S. Z., et al. "Financial toxicity in lung cancer: an assessment of magnitude, perception, and impact on quality of life." *Annals of Oncology* 31.1 (2020): 96-102.
- [3] 김선제 & 사공진. "재난적 의료비와 삶의 질 결정요인에 관한 패널분석." *보건경제와 정책연구 (구 보건경제연구)* 23.4 (2017): 1-22.
- [4] 박진영, 정기택 & 김용민. "중·고령 가구의 과부담 의료비 발생의 결정요인에 관한 패널연구." *보건행정학회지* 24.1 (2014): 56-70.
- [5] Capgemini. (2014). Future value chain 2022. <http://www.theconsumergoodsforum.com/>
- [6] E-daily. (2014). Korea's insurance consumer confidence is the lowest in the world
- [7] H. J. Sim, M. Kim, and H. Choe. (2018). "소비자의 보험소비만족에 관한 융합연구: 자기결정성 이론을 중심으로," *한국융합학회논문지*, vol. 9, no. 5, pp. 157-169
- [8] Kim, Byung-June, and Bong-Gyu Jang. "Old-Age Inequality with Longevity Extension from Preventive Healthcare." Available at SSRN (2020).
- [9] Kuhn, Michael, et al. "Externalities in a life cycle model with endogenous survival." *Journal of Mathematical Economics* 47.4-5 (2011): 627-641.
- [10] Park, Ji-Soo, et al. "Implementation of rule based insurance product recommend and design system using fuzzy inference." *The Journal of Society for e-Business Studies* 12.1 (2007): 99-122.
- [11] Mitra, Sanghamitra, Nilendra Chaudhari, and Bipin Patwardhan. "Leveraging hybrid recommendation system in insurance domain." *International Journal of Engineering and Computer Science* 3.10 (2014).
- [12] Abbas, Assad, et al. "A cloud based health insurance plan recommendation system: A user centered approach." *Future Generation Computer Systems* 43 (2015): 99-109.
- [13] Qazi, Maleeha, et al. "An insurance recommendation system using bayesian networks." *Proceedings of the Eleventh ACM Conference on Recommender Systems*. 2017.
- [14] Y. Boykov and G. Funka-Lea, "Graph Cuts and Efficient N-D Image Segmentation," *IJCV* 70(2), 109-131, 2006.
- [15] Image processing, analysis, and machine vision, Sonka et al., Thomson, 2008.
- [16] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, Aug. 2000.
- [17] Bello, Irwan, et al. "Neural combinatorial optimization with reinforcement learning." *arXiv preprint arXiv:1611.09940* (2016).