

한국사물인터넷학회 종합학술대회

저자일*, 저자이*, 저자삼^{○(발표자 표기)}

*한국컴퓨터대학교 전산학과

[○]한국디지털대학교 컴퓨터공학과

e-mail:{author1, author3}@hcu.ac.kr*, author2@hdu.ac.kr[○]

KIOTS Conference of The Korea IoT Society

Ja-II Jeo*, Ja-E Jeo*, Ja-Sam Jeo[○]

*Dept. of Computer Science, Korea-Computer University

[○]Dept. of Computer Engineering, Korea-Digital University

기초자료를 제공하는 것을 목표로 한다.

I. 서론

현대 사회에서 흡연과 고위험 음주는 국민 건강을 위협하는 주요 위험 요인으로 꾸준히 지목되고 있다. 이러한 건강 위험 행동은 각종 만성질환, 심뇌혈관 질환 및 암 발생률을 높이는 직접적인 원인이 된다. 따라서 개인의 건강 증진을 위해서는 흡연 및 음주 행태를 조기에 발견하고 효과적으로 중재하기 위한 노력이 필수적이다.[1]

기존의 관련 연구들은 주로 특정 질병과의 인과관계를 규명하거나, 특정 인구 집단의 흡연 및 음주 실태를 조사하는 데 중점을 두어 왔다.[2] 최근에는 방대한 건강 데이터 속에서 복합적인 패턴을 발견하고, 이를 기반으로 잠재적 위험군을 예측하려는 데이터 기반의 연구가 주목받고 있다. 특히, 혈액검사와 같은 침습적 방법을 제외하고 신체 계측, 혈압, 생활 습관 설문 등 비침습적 데이터만으로 개인의 건강 위험 행동을 예측하는 것은 보건의료 시스템에서 더욱 간편하고 효율적인 선별 도구로 활용될 수 있을 것이다.[3]

이에 본 연구는 대한민국 국민의 건강 및 영양 상태를 대표하는 국민건강영양조사(KNHANES)의 공개 데이터를 활용한다. 의사결정트리 기반의 머신러닝 모델을 이용하여 개인의 신체 계측 정보, 혈압, 생활 습관 등 비침습적 건강 지표를 복합적으로 분석하고, 이를 통해 현재 흡연 및 고위험 음주 여부를 예측하는 모델을 개발 및 제안한다.[3] 본 연구는 데이터 기반 예측 모델을 통해 건강 위험 행동의 주요 예측 인자를 규명하고, 개인 맞춤형 건강 관리 및 예방 정책 수립을 위한

II. 관련 연구

1. 국내외 연구동향

1.1 국내 동향

흡연과 음주는 주요 공중 보건 문제로서, 이와 관련된 요인을 분석하려는 연구는 꾸준히 수행되어 왔다. 초기 연구들은 주로 국민건강영양조사(KNHANES)와 같은 대규모 데이터를 활용하여 전통적인 통계 기법으로 건강 위험 행동의 실태를 파악하거나 다른 요인과의 관련성을 규명하는 데 중점을 두었다. 예를 들어, 한 연구에서는 KNHANES 원시자료를 이용하여 흡연 및 음주 경험과 안전의식 간의 통계적 유의성을 카이제곱 검정으로 분석하였으며[3], 다른 연구에서는 흡연, 고위험 음주, 신체활동 부족과 같은 복합적인 건강위험행동이 스트레스 및 우울감과 어떤 관련이 있는지 분석했다[1]. 이러한 연구들은 특정 변수들 간의 관계를 이해하는 데 중요한 기초를 제공한다.

최근 대규모 건강 데이터의 활용성이 높아짐에 따라, 데이터마이닝 및 머신러닝 기법을 적용하여 특정 건강 상태나 질병의 위험군을 예측하려는 연구가 활발히 진행되고 있다. 한 예로, KNHANES 데이터를 활용하고 의사결정나무와 같은 데이터마이닝 기법을 적용하여 대사증후군 유병률을 예측하는 모델이 성공적으로 개발되었다[4]. 이 연구는 KNHANES 데이터가 머신러

녕 기반의 예측 모델링에 효과적인 자료원임을 입증하였다[4].

더 나아가, 이러한 예측 방법론은 질병뿐만 아니라 건강 '행동' 자체의 위험 요인을 규명하는 데에도 적용되고 있다. KNHANES 데이터를 활용하여 머신러닝으로 건강 위험 행동을 예측한 연구로는 알코올 사용 장애[2]와 흡연 상태[5]에 대한 예측 연구가 대표적이다. 한 연구에서는 의사결정나무 분석을 적용하여 성인의 알코올 사용 장애 위험 요인을 예측하였으며[2], 다른 연구에서는 여러 머신러닝 알고리즘의 성능 비교를 통해 흡연 상태를 효과적으로 예측할 수 있음을 보였다[5].

이처럼 선행 연구들은 KNHANES 데이터를 활용한 머신러닝 예측 모델의 유효성을 보여주었으나, 대부분이 대사증후군과 같은 특정 질병 예측에 집중되거나[4], 알코올 사용 장애[2] 또는 흡연[5]과 같이 단일 건강 행동에 대한 예측에 머물러 있다. 흡연과 고위험 음주라는 두 가지 핵심적인 건강 위험 행동을 비침습적 데이터만으로 동시에 예측하고, 각 행동의 주요 예측 인자를 비교 분석하려는 통합적인 시도를 하여 비침습적 건강 지표를 활용하여 흡연 및 고위험 음주 여부를 예측하는 통합 모델을 개발하고자 한다.

decision tree analysis," The Journal of Korean Society for School & Community Health Education, 24(1), pp. 47-59, 2023.

- [3] J-O. Kim, H-M. Oh, and J-H. Kim, "A Study on Smoking, Drinking and Safety Awareness using Raw Data in the 8th year of the National Health and Nutrition Examination Survey (2019)," Journal of Digital Policy, 1(1), pp. 27-37, 2022.
- [4] H-K. Kim, K-H. Choi, S-W. Lim, and H-S. Rhee, "Development of Prediction Model for Prevalence of Metabolic Syndrome Using Data Mining: Korea National Health and Nutrition Examination Study," Journal of Digital Convergence, 14(2), pp. 325-332, 2016.
- [5] S. M. Lee, H. Kim, and Y. H. Lee, "A Comparison of Machine Learning Algorithms for Predicting Smoking Status in Korean Adults: Using Korea National Health and Nutrition Examination Survey Data," Journal of the Korean Data & Information Science Society, 33(1), pp. 67-78, 2022.

III. 제안 방법

한국사물인터넷학회와 사물인터넷 관련 산업 발전을 위하여 본 논문에서 제안하고자 하는 내용은 다음과 같다. 이 방법은...

IV. 결론

한국사물인터넷학회와 사물인터넷 관련 산업 발전을 위하여 본 논문에서 제안하는 방법을 이용하면 다음과 같은 결과를 얻을 수 있다. 먼저...

참고문헌

- [1] E-J. Park, J. Jun, and N-S. Kim, "The Association of Multiple Risky Health Behaviors with Self-reported Poor Health, Stress, and Depressive Symptom," Health and Social Welfare Review, 35(1), pp. 136-157, 2015.
- [2] M. Y. Kwon and J. I. Kim, "Risk factors of alcohol use disorder in Korean adults based on the