



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원 저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리와 책임은 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)



순환신경망(RNN) 기법을 통한
고령인구의 일상생활수행능력(ADL)
예측모델 개발

연세대학교 보건대학원

보건정보통계학과 보건정보통계전공

이승미

순환신경망(RNN) 기법을 통한
고령인구의 일상생활수행능력(ADL) 예측모델 개발

지도 하민진 교수

이 논문을 보건학 석사학위 논문으로 제출함

2025년 06월 16일

연세대학교 보건대학원

보건정보통계학과 보건정보통계전공

이승미



이승미의 보건학 석사학위 논문을 인준함.

심사위원 하 민 진 ①

심사위원 박 소희 ①

심사위원 최 동우 ①

연세대학교 보건대학원

2025년 06월 16일

감사의 말씀

본 논문을 완성하기까지 많은 분들의 아낌없는 지도와 격려, 그리고 따뜻한 응원이 있었습니다. 이 자리를 빌어 진심으로 감사의 마음을 전하고자 합니다.

연구의 시작부터 끝까지 아낌없는 지도와 조언을 해주신 하민진 교수님께 깊은 감사의 말씀을 드립니다. 연구 과정 전반에 걸쳐 세심하게 지도해주신 덕분에 부족한 점을 채우고 한 단계 성장할 수 있었습니다.

아울러, 바쁘신 와중에도 소중한 시간을 내어 논문 심사와 조언을 아끼지 않으신 부심 박소희 교수님, 최동우 박사님께도 진심으로 감사드립니다. 귀중한 조언과 따뜻한 격려 덕분에 연구를 더욱 심도 있게 발전시킬 수 있었습니다.

대학원 생활을 함께하며 서로에게 큰 힘이 되어준 동기 여러분께도 고마운 마음을 전합니다. 업무와 학업을 병행하며 모두 힘든 시기가 있었지만 함께 고민하고 격려해주며 성장할 수 있었던 소중한 시간이었습니다.

바쁜 업무 중에도 연구와 학업을 병행할 수 있도록 배려해주신 팀장님과 팀원 분들께도 감사의 말씀을 드립니다. 따뜻한 응원과 이해 덕분에 무사히 논문을 마칠 수 있었습니다.

간호학이라는 익숙한 분야를 떠나 통계라는 새로운 영역에 도전할 수 있었던 것은 무엇보다도 저를 아낌없이 응원해주신 부모님의 든든한 지지 덕분이었습니다. 언제나 한결같은 사랑과 믿음으로 저를 격려해 주셨고, 힘든 순간마다 포기하지 않고 다시 일어설 수 있도록 큰 힘이 되어주셨습니다. 이 자리 를 빌려 저의 성장과 도전을 항상 곁에서 지켜봐 주시고, 묵묵히 응원해주신 부모님께 진심으로 깊은 감사를 드립니다. 앞으로도 가족의 사랑과 믿음을 잊지 않고, 받은 응원에 보답할 수 있도록 더욱 노력하겠습니다.

이 논문은 저 혼자만의 힘으로 완성된 것이 아니라, 많은 분들의 관심과 사랑, 그리고 지지 덕분임을 다시 한 번 마음 깊이 새깁니다. 모든 분들께 진심 으로 감사드립니다.

2025년 6월

이승미 올림

차 례

I . 서론	1
1. 연구 배경 및 연구 필요성	1
2. 연구 목적	5
II . 연구 방법	6
1. 연구자료 및 연구 대상	6
2. 연구 변수	8
2.1 종속변수	8
2.2 독립변수	10
3. 분석 방법	14
3.1 혼합효과 로지스틱 회귀모형	16
3.2 RNN 모델	19
3.3 성능 지표	25
III . 연구 결과	26
1. 연구대상자의 일반적 특성	26
2. 혼합효과 로지스틱 회귀모형의 결과	32
3. 구축된 모형의 성능 비교	38



IV. 고찰	41
1. 연구 결과에 대한 고찰	41
2. 연구의 제한점과 의의	43
2.1. 연구의 제한점	43
2.2. 연구의 의의	44
V. 결론	45
참고문헌	46
ABSTRACT	53

표 차 례

표 1. Questions of K-ADL	9
표 2. Variables used in models	12
표 3. 연구대상자의 일반적 특성 (5차 조사 시점 기준)	30
표 4. 공분산 행렬 구조의 AIC 값 비교 (설명 모델 분석군)	32
표 5. 공분산 행렬 구조의 AIC 값 비교 (예측 모델 분석군)	33
표 6. 혼합 효과 로지스틱 회귀모형의 결과	35
표 7. 예측 모델의 성능 비교	38

그 림 차 례

그림 1. 장래인구추계 - 고령인구(65세 이상) 비중	1
그림 2. 인구피라미드	2
그림 3. Flow chart of subject collection	7
그림 4. Overall analysis process	15
그림 5. ROC Curve	40

국 문 요 약

순환신경망(RNN) 기법을 통한 고령인구의 일상생활수행능력(ADL) 예측모델 개발

본 연구는 고령화폐널 5~9차 자료를 활용하여, 고령인구의 일상생활수행능력(Activities of Daily Living, ADL) 예측을 위한 통계적 및 기계학습 기반 예측모형을 개발하고, 그 성능을 비교하였다. 기존 고령인구 연구에서는 주로 전통적 통계기법인 로지스틱 회귀분석이나 혼합효과 회귀모형을 활용해 왔었다. 이에 본 연구는 시계열 데이터의 장기 의존성 문제를 해결할 수 있는 딥러닝 기반 순환신경망(RNN) 기법(LSTM, GRU)을 적용하여, 고령인구의 ADL 변화 예측에 대한 새로운 접근법을 제시하고자 하였다.

연구대상은 2014년 5차 조사 기준 만 65세 이상이면서 일상생활수행능력(ADL) 점수가 0점(완전자립군)인 고령인구로, 이후 2022년 9차 조사까지 추적된 대상자를 설명 모델 분석군에 포함하였다. 5~9차 조사 중 1번이라도 탈락된 대상자는 제외하고 예측 모델 분석군으로 선정하여 예측 모델 성능 비교에 사용하였다. 주요 독립변수는 성별, 연령, 학력, 가구원 수, 결혼 여부, 주관적 건강상태, 음주 및 흡연 여부, 노동 여부, 가구 총 자산, 고혈압·당뇨 진단, 보청기·틀니 사용 여부 등으로 선정하였다. 예측모형은 혼합효과 로지스틱 회귀모형과 LSTM, GRU 기반의 순환신경망 모델로 구축하였으며, 각 모델의 성능은 AUC, 정확도, 오분류율로 평가하였다.

연구 결과, 혼합효과 로지스틱 회귀모형 결과 독립 변수 중 연령, 주관적 건강상태, 현재 노동 여부, 음주 여부, 보청기 사용 여부 등이 종속변수인 ADL 악화와 유의한 관련성을 보였다. 혼합효과 로지스틱 회귀모형의 AUC는 0.6161, LSTM 모델은 0.7395, GRU 모델은 0.7516으로, 딥러닝 기반 순환신경망 모델이 전통적 통계모형에 비해 우수한 예측력을 보였다. 특히 GRU 모델이 가장 높은 AUC 값을 나타냈으며, 정확도와 오분류율 역시 LSTM, GRU 모델이 더 양호하였다.

결론적으로 본 연구는 반복측정된 고령자 건강자료의 예측에 있어 딥러닝 기반 시계열 모델의 활용 가능성을 실증적으로 제시하였다. 이는 고령인구의 신체기능 저하 예측 및 맞춤형 건강관리 정책 수립에 기여할 수 있는 근거를 제공하며, 향후 일상생활수행능력(ADL) 회복 예측, 해석 가능한 딥러닝 모델 개발 등 다양한 후속 연구의 필요성을 시사한다.

핵심어 : 일상생활수행능력, 혼합효과 로지스틱 회귀모형, 순환신경망 모델, LSTM 모델, GRU 모델

I . 서론

1. 연구 배경 및 연구 필요성

세계적으로 고령인구의 비율은 급속도로 증가하고 있으며 이와 더불어 고령인구의 건강 상태에 대한 관심도 높아졌다. 우리나라에서도 고령인구는 지속적으로 증가하고 있는 추세로 고령화 사회(Ageing society)를 지나서 고령사회(Aged society)로 분류된다 (기획재정부, 2020). 통계청에서 2024년 발표한 우리나라의 고령인구가 우리나라 전체 인구에서 차지하는 비율은 19.2%이다. 고령인구는 지속적으로 증가하고 있으며 2025년에 20%, 2036년에 30%, 2050년에 40% 이상이 될 것으로 예상된다 (그림 1) (통계청, 2024).

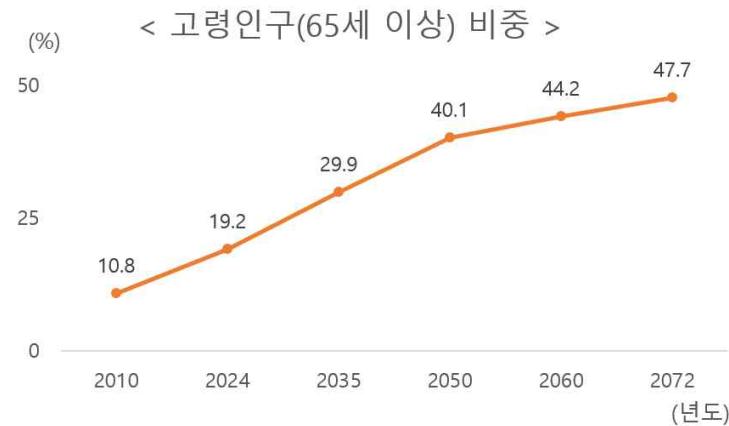


그림 1. 장래인구추계 - 고령인구(65세 이상) 비중 (통계청, 2024)

현재 2022년 인구피라미드는 30대 ~ 50대가 많은 항아리 형이나, 2030년 이후부터 60세 이상이 점차 많아지는 역삼각형 구조로 변화하리라 예상된다. 이를 통해 2025년 우리나라는 65세 이상인 고령인구가 전체 인구에서 차지하는 비율이 20% 이상에 해당하는 초고령 사회(Post-aged society)의 진입이 얼마 남지 않았다는 것을 알 수 있다 (그림 2).

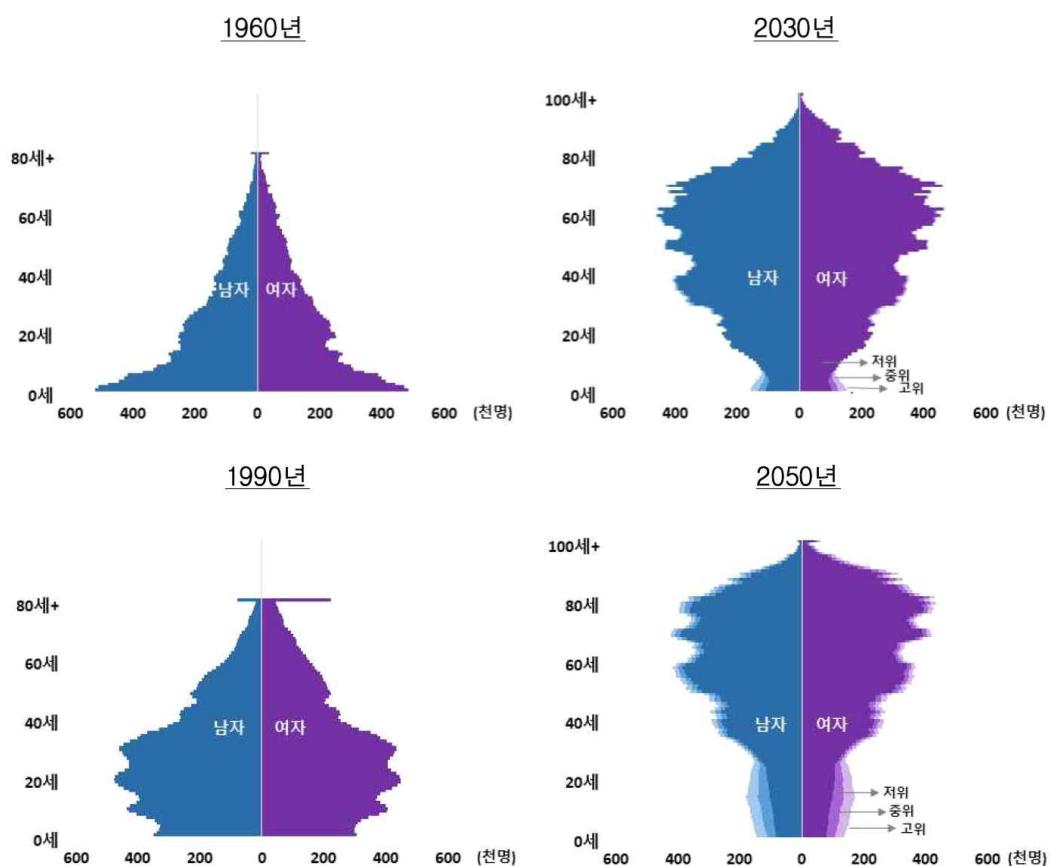


그림 2. 인구피라미드 (통계청, 2024)

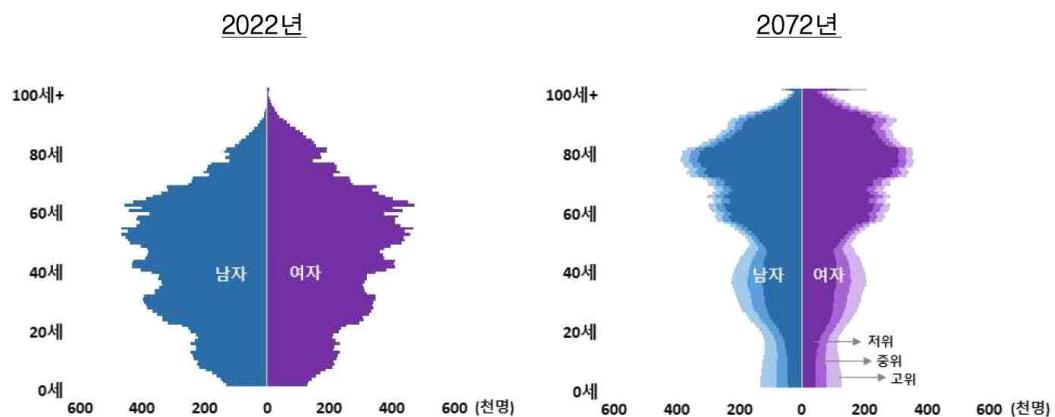


그림 2. 인구피라미드 (통계청, 2024) (계속)

이러한 인구의 노령화는 사회의 다양한 문제를 초래한다. 예를 들어 고령에 의한 건강 쇠퇴는 그 개인의 건강 문제이기도 하지만, 의료 이용량을 증가시키고 사회보장제도의 부담으로 이어지는 등으로 국가적 차원의 문제로 이어질 수 있다 (이소영, 2024). 국민건강증진종합계획(HP2030) 중 5차 계획의 노인 건강 세부 과제내용에 노인의 기능상태 유지와 노인의 장애예방을 위한 지원 항목인 노인성 질환 치료 및 관리 항목이 있을 정도로 노인들의 기능상태 유지는 중요한 과제이다 (보건복지부, 한국건강증진개발원, 2021). 일반적으로 노인의 각종 기능은 청년기 및 장년기에 비해 60에서 70% 정도로 떨어지고, 운동조절 능력 또한 50%대로 떨어진다 (홍갑표, 2003). 이와 같은 결과들로 노인의 각종 기능 상태 평가가 중요하다는 점을 알 수 있다. 세계보건기구 (World Health Organization, WHO)에서는 ‘건강한 노화(Healthy ageing)’을 ‘노년 시기의 웰빙(well-being)이 가능하도록 하는 기능적인 능력을 개발, 유지하는 과정’으로 정의함으로써 광의적 관점에 중점을 두었다 (김유진, 2015).

노인의 신체기능 상태의 평가로 일상생활수행능력 평가를 사용할 수 있다. 이러한 기능 상태 평가는 현재의 기능 상태를 파악하는 데 활용될 뿐 아니라, 향후 장애 발생 가능성까지 예측할 수 있는 잠재력을 지닌다 (이소영, 2024). 이를 통해 고령인구의 자립생활 영위 가능성에 대한 판단이 가능하다 (박주성, 2005). 본 연구에서는 일상생활수행능력(ADL) 예측모형을 개발하여 고령인구의 신체기능 상태 경향을 예측하고자 한다. 2023년 체계적 문헌고찰(Das et al., 2023)에 따르면, 고령자 ADL 예측 연구의 82%가 로지스틱 회귀 또는 랜덤 포레스트를 사용했으며, RNN 계열 모델 적용 사례는 5% 미만으로 확인되었다. 이를 통해 시계열 데이터 처리 능력을 갖춘 딥러닝 접근법이 고령 인구 연구에서 아직 충분히 탐구되지 않았음을 알 수 있다.

본 연구에서는 시계열 데이터를 사용하여 예측하기 위해 기계학습 방법을 적용하였다. 기계학습 방법 중 신경망 모형은 여러 시점에 측정된 값들을 모두 사용하여 고령인구의 일상생활수행능력변화 예측에 적합하다. 이에 따라 신경망 기반 모형과 전통적인 방법인 혼합 효과 회귀모형으로 일상생활수행능력 예측 모형을 개발하여 성능을 비교하고자 한다.

2. 연구목적

본 연구의 목적은 고령화패널 5~9차 자료를 이용하여 노화지표인 일상생활 수행능력(ADL, Activities of Daily Living) 예측모형을 혼합효과 로지스틱 회귀모형(Mixed effects Logistic regression)과 순환신경망 모델(RNN, Recurrent Neural Network)을 각각 개발하여 성능을 비교하는 것이다.

본 연구의 세부목적은 다음과 같다.

첫째, 2014년에 시행한 5차 조사에 참여하고 5차 조사 시 연령이 만 65세 이상인 고령인구를 설명 모델 분석군으로 설정한다. 이후 설명 모델 분석군에서 중도 탈락된 대상자를 제외하여 예측 모델 분석군으로 설정한다. 설명 모델 분석군과 예측 모델 분석군의 일반적인 특성을 파악한다.

둘째, 설명 모델 분석군에 혼합효과 로지스틱 회귀모형을 적용하여 선정한 독립 변수들이 고령인구의 일상생활수행능력으로 분류되는 완전자립군, 도움 필요군에 미치는 영향을 살펴본다.

셋째, 예측 모델 분석군에 예측을 보는 모델로 혼합효과 로지스틱 회귀모형과 기계학습 기반 RNN 모델 중 LSTM(Long short term memory) model, GRU(Gated Recurrent Unit) model을 적용하여 어떤 모델이 좋은 성능을 보이는지 확인해본다.

II . 연구 방법

1. 연구 자료 및 연구 대상

고령화연구패널 조사는 2006년 제주도를 제외한 우리나라 전국의 만 45세 이상인 10,254명을 원표본으로 하고 5차 시기 만 45세 이상인 920명 표본을 충원하여 신규 패널로 구분한 후 2년마다 추적조사를 시행하여 현재까지 9차 추적조사가 완료된 조사이다. 9차 조사에 이르는 동안 기준 패널의 표본 유지율은 77.2%를 기록하고 있다 (한국고용정보원, 2024).

고령화연구패널 조사의 1차 조사가 2006년에 만 45세 이상인 대상자들로 구성되었고 본 연구에서는 만 65세 이상인 대상자들의 일상생활수행능력을 살펴보려고 하기 때문에 첫 조사 이후 2년마다 조사가 시행되어 더 많은 대상자들을 분석에 사용하기 위해 2014년에 시행한 5차 조사부터 참여한 대상자들을 본 연구에서 대상자로 선정하였다.

본 연구에서는 2014년에 시행한 5차 조사에 참여한 대상자 7,949명 중 5차 조사 시점 기준으로 연령이 만 65세 이상인 대상자 4,259명을 추출하였다. 그 후 5차 조사 시점 기준으로 일상생활수행능력(ADL) 점수가 0점인 대상자만 추출하여 설명 모델 분석군 3,935명을 선정하여 이미 일상생활수행능력에 제한이 있던 대상자를 제거하며 혼란 변수를 최소화하였다. 이후 5~9차 조사 시점에서 탈락된 대상자 1,398명을 제거하여 예측 모델 분석군으로 2,537명을 선정하였다.

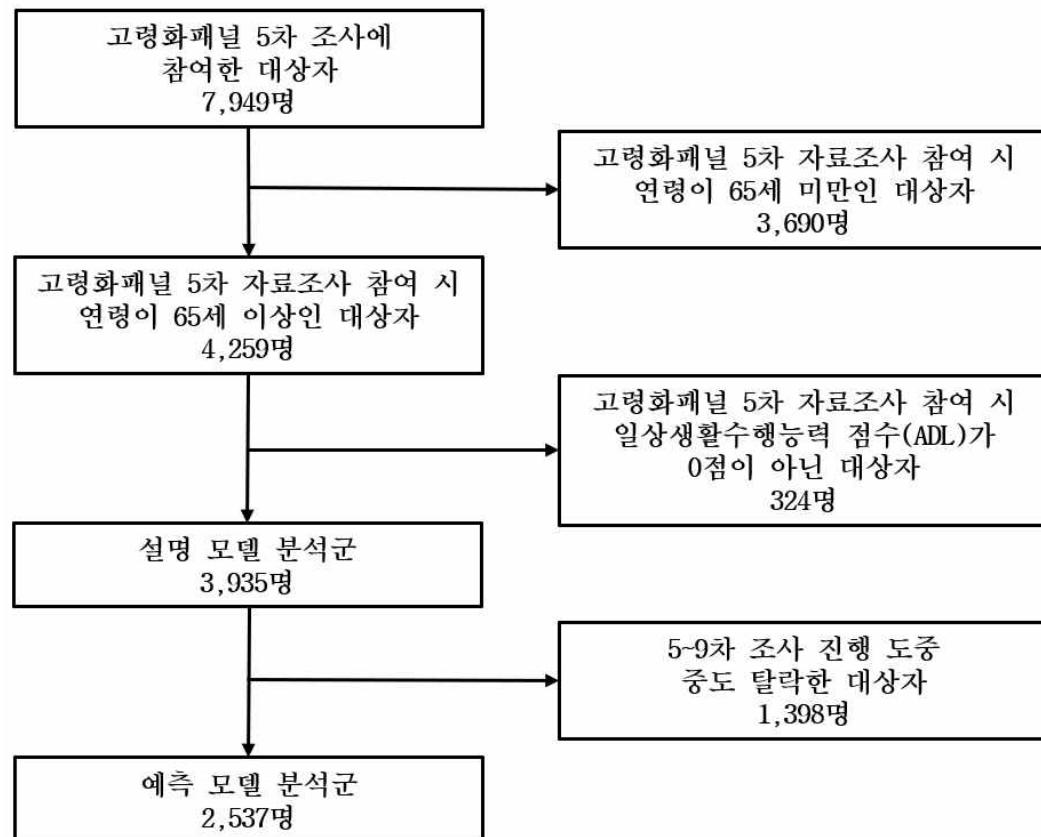


그림 3. Flow chart of subject collection

2. 연구 변수

본 연구는 2년 주기로 반복 측정된 고령화폐널 연구조사 자료를 사용하여, 일상생활수행능력을 예측하는 모델을 개발하기 위해 변수들을 다음과 같이 선정하였다 (표 2).

2.1. 종속변수

본 연구의 종속변수는 일상생활수행능력 점수이다. 고령화폐널 연구조사에서 사용한 일상생활수행능력 측정도구는 2002년에 대한노인병학회 노인기능평가연구회에서 한국 노인의 기능을 평가하기 위해 개발한 측정도구인 한국형 일상생활수행능력 측정도구(K-ADL)을 사용하였다 (박수지, 박병선, 2017). 한국형 일상생활수행능력 측정도구의 계산된 신뢰도는 Cronbach's alpha를 사용하여 계산된 값으로 0.937이었다 (원장원 외, 2002).

고령화폐널 연구조사에서는 한국형 일상생활수행능력 측정도구를 활용한 7 가지 문항(옷 갈아입기, 세수/양치/머리감기, 목욕/샤워하기, 식사하기, 방 밖으로 나가기, 화장실 이용하기, 대소변 조절하기)으로 평가하였다 (표 1).

부분적인 도움 또는 전적인 도움이 필요한 경우는 '1'의 값으로, 도움이 필요하지 않으면 '0'의 값으로 각 항목을 계산하여 0~7점까지의 값을 가지게 된다. 0점에 가까울수록 일상생활 수행 시 도움이 필요하지 않다는 의미이며 7점에 가까울수록 일상생활 수행 시 도움이 필요하다는 의미로 해석한다.

본 연구에서는 한국형 일상생활수행능력 점수가 0일 ‘완전자립군’과 1~7 점인 ‘도움필요군(부분 or 전적)’으로 대상자를 구분하여 분석하였다.

표 1. Questions of K-ADL

Questions	Answer	ADL Category
최근 일주일 동안의 활동을 기준으로 도움 필요 없음		완전자립군
말씀해 주시고, 지금은 잠시 아프시거나		
나 다치셔서 도움을 받으시지만, 앞으로 3개월 이내에 고쳐질 것으로 예상되는 것은 도움이 필요 없는 것으로 말씀해 주십시오.	부분적인 도움이 필요함 전적으로 도움이 필요함	도움필요군 (부분 or 전적)
1) 옷 갈아입기 2) 세수/양치/머리감기 3) 목욕/샤워하기 4) 식사하기 5) 방 밖으로 나가기 6) 화장실 이용하기 7) 대소변 조절하기		

2.2. 독립변수

본 연구에서는 성별, 연령, 최종 학력, 가구원 수, 결혼 여부, 주관적 건강 상태, 음주 여부, 흡연 여부, 현재 노동 여부, 가구 총 자산, 고혈압 진단 여부, 당뇨 진단 여부, 보청기 사용 여부, 틀니 사용 여부, 병의원 방문 횟수를 독립 변수로 선정하였다.

연령은 총 3개의 군인 ‘65세 이상 75세 미만’인 군과 ‘75세 이상 85세 미만’인 군, 그리고 ‘85세 이상’인 군으로 분류하여 범주형 변수로 사용하였다. 최종 학력은 ‘초등학교 졸업 이하’와 ‘중학교 졸업’, 그리고 ‘고등학교 졸업’과 ‘고등학교 이상 졸업’으로 분류하였다. 가구원 수는 현재 함께 거주하고 있는 가구원의 수를 의미하며 ‘1명’과 ‘2명과, ‘3명’, 그리고 ‘4명 이상’으로 분류하여 범주형 변수로 사용하였다. 결혼 여부는 ‘미혼’과 ‘혼인’과 ‘별거/이혼/사별’로 분류하였다. 주관적 건강상태는 주관적 평가에 따라 ‘좋음’과 ‘보통’과 ‘나쁨’으로 분류하였다. 음주 여부는 평생 음주 경험이 없는 집단인 ‘비음주자’, 과거에는 음주 경험이 있지만 현재 음주하지 않는 집단인 ‘과거음주자’, 과거 음주 경험 유무와 관계없이 현재 음주하는 집단인 ‘현재음주자’로 분류하였다. 흡연 여부도 동일하게 평생 흡연 경험이 없는 집단인 ‘비흡연자’, 과거에는 흡연 경험이 있지만 현재 흡연하지 않는 집단인 ‘과거흡연자’, 과거 흡연 경험 유무와 관계없이 현재 흡연하는 집단인 ‘현재흡연자’로 분류하였다.

가구 총 소득은 가구 균등화 소득 산정 방법(가구 균등화 소득= $\frac{\text{가구총소득}}{\sqrt{\text{가구구성원수}}}$)로 보정하여 사용하였다. 대상자들의 제1사분위수와 제3사분위수를 구하여 707만원, 1905만원을 기준으로 구분하여 가구 총 소득이 707만원 미만인 경우 ‘저소득’, 707만 원 이상 1905만원 미만인 경우 ‘중산층’, 1905만 원 이상인 경우 ‘고소득’으로 분류하여 범주형 변수로 사용하였다. 고혈압 진단 여부는 ‘진단받지 않음’, ‘고혈압’으로 분류하였다. 당뇨 진단 여부는 ‘진단받지 않음’, ‘당뇨’로 분류하였다. 보청기 사용 여부는 ‘사용하지 않음’, ‘사용함’으로 분류하였다. 틀니 사용 여부는 ‘사용하지 않음’, ‘사용함’으로 분류하였다.

표 2. Variables used in models

역할	변수명	분류
종속변수	일상생활수행능력(ADL)	1) 완전자립군 (0점) 2) 도움필요군 (1~7점)
독립변수	성별	1) 남자 2) 여자
	연령 그룹	1) 65세 이상 75세 미만 2) 75세 이상 85세 미만 3) 85세 이상
	최종 학력	1) 초등학교 졸업 이하 2) 중학교 졸업 3) 고등학교 졸업 4) 대학교 이상 졸업
	가구원 수 그룹	1) 1명 2) 2명 3) 3명 4) 4명 이상
	결혼 여부	1) 미혼 2) 혼인 중 3) 별거/이혼/사별

표 2. Variables used in models (계속)

역할	변수명	분류 및 단위
독립변수	주관적 건강상태	1) 좋음 2) 보통 3) 나쁨
	음주 여부	1) 비음주자 2) 과거음주자 3) 현재음주자
	흡연 여부	1) 비흡연자 2) 과거흡연자 3) 현재흡연자
	현재 노동 여부	1) 노동하지 않음 2) 노동함
	가구 총 자산	1) 저소득 2) 중산층 3) 고소득
	고혈압 진단 여부	1) 진단받지 않음 2) 고혈압
	당뇨 진단 여부	1) 진단받지 않음 2) 당뇨
	보청기 사용 여부	1) 사용하지 않음 2) 사용함
	틀니 사용 여부	1) 사용하지 않음 2) 사용함

3. 분석 방법

본 연구에 사용된 통계분석은 SAS(Statistical Analysis System version 9.4)와 Python (version 3.13) 프로그램을 사용하였다.

본 연구의 분석 과정은 그림 4와 같이 정리하였다. 앞서 연구 대상 부분에서 정리한 설명 모델 분석군, 예측 모델 분석군으로 선정된 대상자들의 일반적인 특성에 대해 살펴보았다. 설명 모델 분석군에 혼합 효과 로지스틱 회귀 모형을 적용하여 선정한 독립 변수들이 고령인구의 일상생활수행능력으로 분류되는 완전자립군, 도움필요군에 미치는 영향을 살펴보았다. 전체 데이터를 무작위로 70%의 Training set, 30%의 Testing set으로 분할하여 RNN 모델에 사용하였으며 재현성을 위해 랜덤 시드(Random seed)를 42로 고정하였다. 예측 모델 분석군을 혼합 효과 로지스틱 회귀모형에 적용한 결과와 Training set, Testing set으로 분할한 예측 모델 분석군에 LSTM 모델, GRU 모델을 적용한 총 결과 총 3가지 예측 모델의 성능을 비교하였다. Model 간 성능 비교를 위한 척도로 정확도(Accuracy), 오분류율(Misclassification rate), AUC(Area Under the Curve)값을 사용하였다.

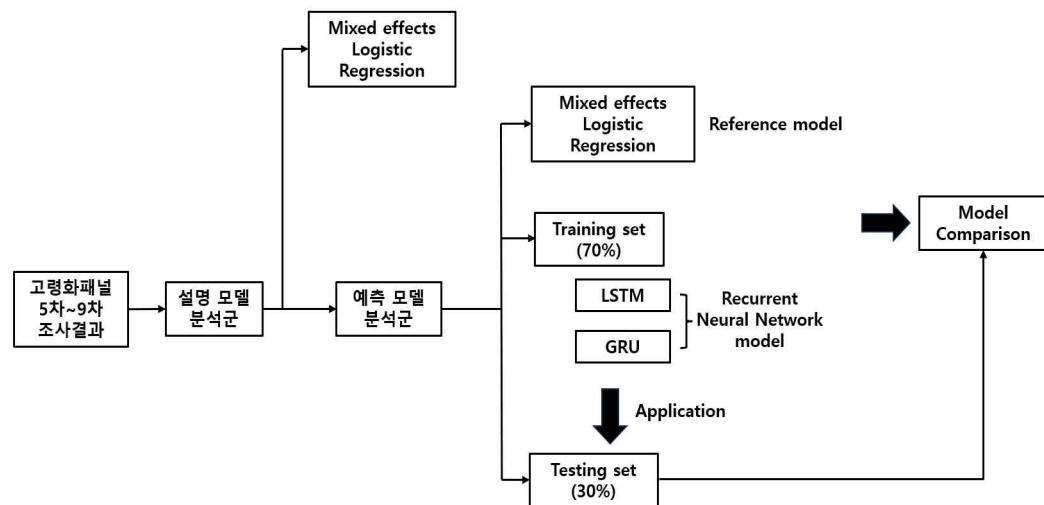


그림 4. Overall analysis process

3.1. 혼합 효과 로지스틱 회귀모형

혼합 효과 로지스틱 회귀모형은 반복 측정된 자료를 이용하여 이진형 종속 변수와 독립 변수 간의 관계를 설명할 때 사용된다. 이 모형은 개인 간 변동을 통제하기 위해 대상자별 임의 효과(Random Effects)을 포함한 일반화 선형 모형(GLMM, Generalized Linear Mixed Model)으로 볼 수 있다. 혼합 효과 로지스틱 회귀모형에서는 종단적 반응과 중도 탈락 메커니즘을 각각 별도로 모델링하면서도 이 두 과정을 임의 효과를 통해 연결한다. 고정효과(Fixed Effects)는 관심 있는 예측변수 즉, 종속변수의 평균적 효과를 추정하는 효과로 본 연구에서는 성별, 연령, 흡연 여부, 음주 여부, 고혈압 진단 여부 등이 포함된다. 임의효과는 관측되지 않은 개인 또는 집단 간 변동성을 모델링하는 효과로 본 연구에서는 시간적 반복 측정과 대상자 개인이 될 수 있다.

본 연구에서는 5차수의 반복측정 자료의 시간적 상관성을 적절히 반영하기 위해 다양한 공분산 행렬 구조를 비교하였다. 분산성분(VC), 복합대칭성(CS), 비구조적(UN), 1차 자기회귀(AR(1)), Toeplitz(TOEP)의 AIC, AICC, BIC 값을 비교하여 주요 공분산 구조를 선택하였다.

분산성분(VC, Variance Components) 구조는 가장 단순한 구조로 계산 효율성이 높으나 개인 간 변동만을 고려하며, 시간적 상관성을 무시하는 구조로 반복측정 자료의 시간적 의존성 반영하기 힘들다는 단점이 있다. 복합대칭성(CS, Compound Symmetry) 구조는 모든 시점 간 상관계수가 동일하고 분산이 균일하다는 가정을 가지며 보통 시간에 무관한 상관성 모델링 연구에 사용된다.

비구조적(UN, Unstructured) 구조는 시점별 분산과 공분산을 자유롭게 추정하는 가장 유연한 구조이지만 추정할 모수가 많아 표본 크기가 작을 경우 과적합의 위험이 있다. 1차 자기회귀(AR(1)) 구조는 시간 간격이 증가함에 따라 상관계수가 지수적으로 감소하는 구조이며 혈압, 혈당 등 시간에 따른 점진적 변화가 예상되는 연구에 주로 사용되는 구조이다. 마지막으로, Toeplitz(TOEP) 구조는 동일한 시간 간격을 갖는 관측치 간 상관계수가 동일한 구조이며 1차 자기회귀(AR(1)) 구조보다는 유연하며 비구조적(UN) 구조보다는 모수의 수가 적은 특징을 가진다 (George et al., 2015).

공분산 구조를 선택하기 위해 비교한 값은 AIC, AICC, BIC 값이다. 먼저, Akaike 정보 기준(Akaike Information Criterion, AIC)은 통계 모델의 적합도를 평가하고 여러 후보 모델 중 가장 적합한 모델을 선택하기 위해 사용되는 지표이다. AIC는 모델의 최대우도(maximum likelihood)와 모델에 포함된 파라미터의 수를 모두 고려하여 계산되며, 모델의 설명력이 높으면서도 과적합을 방지하는 모델을 선호하도록 설계되었다. AIC 값이 낮을수록 데이터에 더 적합한 모델로 간주된다 (Cavanaugh et al., 2019). AICC(Corrected AIC)는 AIC의 단점인 소표본에서의 과적합 경향을 보완하기 위해 개발된 수정 지표로, 표본 크기가 작을 때 추가적인 패널티 항을 더하여 보다 보수적으로 모델을 평가한다. AICC는 표본 크기가 충분히 크면 AIC와 거의 동일해지지만, 표본이 적을 때는 AICC를 사용하는 것이 권장된다 (Cavanaugh et al., 2019). Bayesian 정보 기준(BIC, Bayesian Information Criterion)은 AIC와 유사하지만, 모델 복잡성에 대한 패널티가 더 강하게 적용된다. BIC는 표본 크기가 커질수록 복잡한 모델에 더 큰 패널티를 부과하며, 단순한 모델을 선호하는 경향이 있다 (Cavanaugh et al., 2019). BIC 역시 값이 낮을수록 더 적합한 모델로 간주된다.

AIC, AICC, BIC 모두 여러 후보 모델 간의 상대적 적합도를 비교하는 데 사용되며, 가장 낮은 값을 가진 모델이 최적의 모델로 선택된다. 본 연구에서도 이 3가지의 값을 비교하여 혼합 효과 로지스틱 모형을 적용할 때 공분산 구조를 선택하였다.

혼합 효과 로지스틱 모형은 중도 탈락이 많은 경우, 정보적 탈락 (informative dropout)으로 인한 편향이 발생할 수 있다 (Ten Have et al., 1998). 정보적 탈락이란 탈락 확률이 관찰되지 않은 결과나 잠재적 요인에 의존하는 경우를 의미하며, 이러한 상황에서 단순히 무작위 탈락으로 가정하고 분석을 수행할 경우 모수 추정에 심각한 편향이 발생할 수 있다. 실제로, 정보적 탈락을 고려하지 않은 혼합 효과 로지스틱 회귀모형에서는 회귀계수 추정치가 왜곡되거나 신뢰구간의 커버리지가 저하되는 등 통계적 추론의 정확성이 떨어질 수 있다. 따라서 중도 탈락이 많은 종단 자료에서는 탈락 메커니즘을 명시적으로 모델에 포함시키는 등 적절한 통계적 접근이 필요하다 (Ten Have et al., 1998).

본 연구에서는 LSTM, GRU 모델과 혼합 효과 로지스틱 회귀모형의 성능 지표 (AUC, 정확도, 오분류율)를 비교하여 모델 간 예측 성능 차이를 규명하였다.

3.2. RNN 모델

본 연구에서는 시계열 데이터의 특성을 반영하기 위해 RNN 모델 기반의 LSTM 및 GRU 모델을 적용하였다. 기존의 다층 퍼셉트론 등 신경망(Neural Network) 모델은 입력 간의 시계열적 의존성을 고려하지 않는다 (유건령, 2020). 반면, 기존의 신경망 모형과는 다르게 독립적이고 순차적인 정보를 처리하기 위해 개발된 RNN 모델은 독립적이고 순차적인 정보 처리에 특화되어 이전 시점의 은닉 상태 정보를 현재 시점에 반영함으로써 시간적 또는 순차적 데이터의 패턴을 효과적으로 학습한다 (이수지, 2016). RNN 모델은 피드백 구조를 통해 이전 상태의 정보를 내부적으로 저장하고, 이를 다음 단계의 입력과 결합하여 새로운 은닉 상태를 계산한다. 이때 은닉 상태(hidden state)는 시간에 따라 순차적으로 갱신되며, 각 시점의 은닉 상태는 이전 시점의 은닉 상태와 현재 입력의 함수로 정의된다. 학습 과정에서는 역전파 알고리즘을 시간에 따라 확장한 ‘시간에 따른 역전파(Backpropagation Through Time, BPTT)’ 방법이 사용된다. 이 방법은 순환신경망 구조를 시퀀스 길이만큼 펼쳐 각 시점의 은닉 상태와 출력을 계산한 후, 전체 시퀀스에 대한 손실을 계산하여 가중치와 편향 등 파라미터를 업데이트하는 방식이다 (김호현, 2017). 시계열 데이터의 마지막 시점부터 첫 시점까지 역방향으로 손실을 계산하며, 각 시점의 손실이 이전 시점의 은닉 상태와 입력에 미치는 영향을 누적적으로 계산하는 체인 룰(chain rule)이 적용된다 (Cho et al., 2014). 이를 통해 RNN 모델은 시퀀스 내에서 시간적으로 떨어진 입력 간의 의존성을 효과적으로 학습할 수 있다(Chung et al., 2014).

한계점으로는, 시퀀스가 길어지면서 어느 단계부터 더 이상 학습이 이루어 지지 않아서 과거의 정보가 소실되는 ‘장기 의존성 문제(the problem of Long-Term Dependencies)’가 발생한다. 장기 의존성 문제는 시퀀스가 길어질수록 앞의 정보가 뒤의 정보에 충분히 전달하지 못하는 현상으로 시간이 지날수록 초기 정보가 소실되어 최종 예측 값에 대한 영향력이 떨어지는 것이다 (김호현, 2017). 이러한 한계를 해결하기 위한 방법으로 LSTM 모델이 개발되었다. LSTM 모델은 구조가 복잡하고 학습 시간이 오래 소요된다는 단점이 있어 이를 보완하여 GRU 모델이 개발되었다 (김호현, 2017).

RNN 모델의 성능은 층 수, 각 층의 유닛 수, Drop-out 비율, 학습률, Batch 크기, Epoch 등 다양한 하이퍼파라미터에 의해 영향을 받는다. LSTM/GRU 층을 여러 개 쌓을수록 더 복잡한 시계열 패턴을 학습할 수 있지만, 과적합의 위험성이 증가한다. 각 LSTM/GRU 층의 유닛 수가 많을수록 더 많은 정보를 학습할 수 있으나 계산량이 늘어나 학습 속도가 지연될 수 있다. Drop-out 비율은 과적합 방지를 위해 일부 뉴런의 출력을 무작위로 0으로 만드는 비율로, 일반적으로 0.05~0.5 사이로 설정된다. 학습률은 가중치 업데이트 크기를 결정하며, 너무 작으면 가중치가 조금씩만 조정되어 학습이 느려질 수 있고, 너무 크면 가중치가 너무 크게 조정되어 불안정한 학습이 발생할 수 있다. Batch 크기는 한 번에 학습하는 데이터 샘플의 수로, 메모리와 학습 속도에 영향을 미친다. Epoch는 전체 훈련 데이터를 몇 번 반복해서 학습할지 결정한다 (Kharel et al., 2024).

탐색 대상 하이퍼파라미터와 범위는 다음과 같이 설정하였다. 순환층 유닛 수는 [64, 128, 256] 중 선택하였고, 드롭아웃 비율은 [0.03, 0.05, 0.1, 0.2], 학습률은 [0.0001, 0.0002, 0.0005, 0.001], 배치 크기는 [32, 64, 80], 에포크 수는 [100, 150, 200] 중 선택하였다. 이를 통해 LSTM 모델, GRU 모델 각각에서 AUC 값이 최대가 되는 하이퍼파라미터 조합을 최종적으로 채택 하였으며, 최적의 하이퍼파라미터를 적용해 전체 훈련 데이터를 재학습한 후 테스트 데이터에서 최종 성능을 평가하였다. 이러한 과정을 통해 모델의 일반화 성능을 극대화하고, 실험의 재현성을 확보하였다.

3.2.1. LSTM 모델

LSTM 모델은 기존의 RNN 모델의 주요 문제인 장기 의존성 문제를 해결하기 위해 고안된 신경망 구조이다. LSTM 모델은 셀 상태와 게이트 구조를 통해 장기 의존성 문제를 해결한다 (Kong et al., 2024).

LSTM 모델은 셀 상태와 입력 게이트, 망각 게이트, 출력 게이트 총 3 가지의 주요 게이트로 구성된다. 입력 게이트(Input Gate)는 현재 시점의 입력과 이전 시점의 은닉 상태를 바탕으로, 새로운 정보 중 셀 상태에 추가할 정보들을 결정한다. 입력 게이트의 출력값이 1에 가까울수록 해당 정보가 많이 반영된다(Han et al., 2020). 망각 게이트(Forget Gate)는 이전 시점의 셀 상태에서 어떤 정보를 삭제할지 결정하여 불필요한 정보에 대한 기억을 제거한다. 망각 게이트의 출력값이 0에 가까울수록 해당 정보는 잊혀지고, 1에 가까울수록 유지된다 (Han et al., 2020). 출력 게이트(Output Gate)는 셀 상태로부터 어느 정도의 정보를 은닉 상태로 내보낼지 결정하여 출력할 정보를 정한다 (Han et al., 2020). 셀 상태(Cell state)는 장기 기억을 저장하는 경로로, 시퀀스 전체를 따라 거의 변형 없이 정보를 전달할 수 있다. 이 셀 상태는 각 게이트를 통해 정보가 추가되거나 삭제될 수 있다. 망각 게이트와 입력 게이트의 결과를 조합하여 셀 상태를 업데이트한다 (Kharel et al., 2024). 은닉 상태(Hidden state)는 LSTM 모델에서 현재 시점까지 입력된 정보를 요약하여 담고 있는 벡터로, 시계열 데이터의 시간적 패턴을 인코딩하는 역할을 한다. 은닉 상태는 LSTM 셀의 각 시점마다 계산되며, 다음 시점의 입력으로 전달될 뿐만 아니라, 필요에 따라 모델의 출력층으로 직접 연결될 수 있다.



본 연구에서의 LSTM 모델에 대한 하이퍼파라미터 탐색 및 선정 과정은 앞선 순환신경망 모델 연구 방법에 준하여 동일하게 적용하였다. 2개의 LSTM 층(각 256, 128개의 유닛 수)과 2개의 Dense 층(각 128, 64개의 유닛 수)으로 구성된 모델을 구축하였다. Drop-out 비율 0.05를 적용하여 과적합을 방지하였으며 초기 학습률을 0.0001로 설정하였다. Batch 크기는 32, Epoch는 150으로 LSTM 모델을 적용하여 예측 모델을 개발하였다.

3.2.2. GRU 모델

GRU 모델은 LSTM 모델의 구조를 단순화하여 빠른 학습 속도를 가지며 LSTM 모델과 비슷한 성능을 보이도록 고안된 model이다 (조경현, 2014). 또한 데이터가 충분하지 않은 경우 과적합의 문제가 발생하는 LSTM 모델의 한계점을 해결하기 위해 고안되었다. LSTM 모델은 3개의 게이트가 존재하였으나 GRU 모델은 업데이트 게이트와 리셋 게이트 총 2개의 게이트만을 사용하며, LSTM 모델 대비 파라미터 수가 적고 학습 속도가 빠르다.

업데이트 게이트(Update Gate)는 이전 상태의 정보를 어느 정도 유지시킬지 결정하는 역할을 한다 (Chung et al., 2014). 리셋 게이트(Reset Gate)는 현재 상태에서 이전 상태에 대한 정보를 어느 정도 잊어야 하는지 결정하는 역할을 한다. LSTM 모델의 셀 상태 대신 은닉 상태를 사용하여 과적합의 문제를 해결 한다 (Chung et al., 2014). 은닉 상태는 GRU 모델에서 각 시점마다 출력되며 현재 시점까지 입력된 정보를 요약한 값이다.

본 연구에서의 GRU 모델에 대한 하이퍼파라미터 탐색 및 선정 과정은 앞선 순환신경망 모델 연구 방법에 준하여 동일하게 적용하였다. 2개의 GRU 층(각 256, 128개의 유닛 수)과 2개의 Dense 층(각 128, 64개의 유닛 수)으로 구성된 모델을 구축하였다. Drop-out 비율 0.03을 적용하여 과적합을 방지하였으며 초기 학습률을 0.0002로 설정하였다. Batch 크기는 80, Epoch는 200으로 GRU 모델을 적용하여 예측 모델을 개발하였다.

3.3. 성능 지표

본 연구에서는 혼합 효과 로지스틱 회귀모형과 RNN 모델인 LSTM 모델과 GRU 모델 사이의 성능 비교를 위해 AUC값, 정확도(Accuracy), 오분류율 (Misclassification rate)을 산출하였다.

정확도는 연구에 사용된 전체 데이터 기준 모델이 올바르게 예측한 비율을 의미하며 전체 관측치 중에서 올바르게 분류된 사례의 비율을 의미한다 (김철립. 2023). 오분류율은 전체 데이터 중 모델이 잘못 예측한 비율을 의미하며 (1-정확도)로 계산할 수 있다.

AUC값은 ROC curve 아래의 면적을 의미하며, 분류 모델의 성능을 평가하는 대표적인 지표 중 하나이다. AUC 값은 0과 1 사이의 값을 가지며, 1에 가까울수록 모델이 양성(positive)과 음성(negative) 클래스를 더 정확하게 구분함을 의미한다. 반면, 0.5에 가까운 값은 모델이 무작위 추측(random guess)과 다를 바 없음을 나타내며, 0에 가까울수록 분류 성능이 낮다는 것을 의미한다 (Anne et al., 2022).

III. 연구 결과

1. 연구대상자의 일반적 특성

설명 모델 분석군에 선정된 대상자는 총 3,935명이었으며 5차 조사 자료 기준으로 연구대상자의 일반적 특성을 살펴보면 다음과 같다 (표 3).

성별을 살펴보면 남자인 대상자는 1,678명으로 전체 대상자의 42.64%였고, 여자는 2,257명으로 전체 대상자의 57.36%로 더 많았다. 연령 그룹을 살펴보면 65세 이상 75세 미만에 해당하는 대상자는 2,135명으로 전체 대상자의 54.26%였고, 75세 이상 85세 미만인 대상자는 1,507명으로 전체 대상자의 38.30%였고, 85세 이상인 대상자는 293명으로 전체 대상자의 7.45%였다. 최종 학력을 살펴보면 초등학교 졸업 이하에 해당하는 대상자는 2,387명으로 전체 대상자의 60.66%였고, 중학교 졸업에 해당하는 대상자는 589명으로 전체 대상자의 14.97%였다. 고등학교 졸업에 해당하는 대상자는 676명으로 전체 대상자의 17.18%였고, 대학교 이상 졸업에 해당하는 대상자는 283명으로 전체 대상자의 7.19%였다. 가구원 수를 살펴보면 가구원 수가 1명인 대상자는 770명으로 전체 대상자의 19.57%였고 가구원 수가 2명인 대상자는 2,292명으로 전체 대상자의 58.25%로 가장 많았다. 가구원 수가 3명인 대상자는 466명으로 전체 대상자의 11.84%였고, 가구원 수가 4명 이상인 대상자는 407명으로 전체 대상자의 10.34%였다.

결혼 여부에서 미혼인 대상자는 11명으로 전체 대상자의 0.28%를 였고, 혼인 중인 대상자는 1,253명으로 전체 대상자의 31.84%였고, 별거/이혼/사별에 해당하는 대상자는 2,671명으로 전체 대상자의 67.88%로 가장 많았다. 주관적 건강상태를 살펴보면 주관적 건강상태가 좋음인 대상자는 633명으로 전체 대상자의 16.09%를 차지하였고, 주관적 건강상태가 보통인 대상자는 1,843명으로 전체 대상자의 46.84%를 차지하였고, 주관적 건강상태가 나쁨인 대상자는 1,459명으로 전체 대상자의 37.08%를 차지하였다. 음주 여부를 살펴보면 비음주자는 2,181명으로 전체 대상자의 55.43%였고, 과거음주자는 680명으로 전체 대상자의 17.28%였고, 현재음주자는 1,074명으로 전체 대상자의 27.29%였다. 흡연 여부를 살펴보면 비흡연자는 2,759명으로 전체 대상자의 70.11%였고, 과거흡연자는 779명으로 전체 대상자의 19.80%였고, 현재흡연자는 397명으로 전체 대상자의 10.09%였다. 현재 노동 여부를 살펴보면 노동하지 않는 대상자는 2,962명으로 전체 대상자의 75.27%였고, 노동하는 대상자는 973명으로 전체 대상자의 24.73%였다. 가구 총 자산을 살펴보면 저소득에 해당하는 대상자는 2,290명으로 전체 대상자의 58.20%였고, 중산층에 해당하는 대상자는 773명으로 전체 대상자의 19.64%였고, 고소득에 해당하는 대상자는 872명으로 전체 대상자의 22.16%였다. 고혈압 진단 여부를 살펴보면 고혈압 진단받지 않은 대상자는 1,861명으로 전체 대상자의 47.29%였고, 고혈압 진단을 받은 대상자는 2,074명으로 전체 대상자의 52.71%였다. 당뇨 진단 여부를 살펴보면 당뇨 진단받지 않은 대상자는 3,040명으로 전체 대상자의 77.26%였고, 당뇨 진단받은 대상자는 895명으로 22.74%였다. 보청기 사용 여부를 살펴보면 보청기를 사용하지 않는 대상자는 3,831명으로 전체 대상자의 97.36%였고, 보청기를 사용하는 대상자는 104명으로 전체 대상자의 2.64%였다. 틀니 사용 여부에서 틀니를 사용하지 않는 대상자는 2,656명으로 전체 대상자의 67.50%였고, 틀니를 사용하는 대상자는 1,279명으로 전체 대상자의 32.50%였다.

예측 모델 분석군으로 선정된 2,537명의 일반적 특성을 살펴보면 다음과 같다 (표 3).

성별을 살펴보면 남자인 대상자는 1,033명으로 전체 대상자의 40.72%였고, 여자는 1,504명으로 전체 대상자의 59.28%로 더 많았다. 연령 그룹을 살펴보면 65세 이상 75세 미만에 해당하는 대상자는 1,656명으로 전체 대상자의 65.27%였고, 75세 이상 85세 미만인 대상자는 815명으로 전체 대상자의 32.12%였고, 85세 이상인 대상자는 66명으로 전체 대상자의 2.60%였다. 최종 학력을 살펴보면 초등학교 졸업 이하에 해당하는 대상자는 1,656명으로 전체 대상자의 65.27%였고, 고 중학교 졸업에 해당하는 대상자는 401명으로 전체 대상자의 15.81%였다. 고등학교 졸업에 해당하는 대상자는 477명으로 전체 대상자의 18.80%였고, 대학교 이상 졸업에 해당하는 대상자는 186명으로 전체 대상자의 7.33%였다. 가구원 수를 살펴보면 가구원 수가 1명인 대상자는 417명으로 전체 대상자의 16.44%였고 가구원 수가 2명인 대상자는 1,581명으로 전체 대상자의 62.32%였다. 가구원 수가 3명인 대상자는 297명으로 전체 대상자의 11.71%였고, 가구원 수가 4명 이상인 대상자는 242명으로 전체 대상자의 9.54%였다. 결혼 여부를 살펴보면 미혼인 대상자는 8명으로 전체 대상자의 0.32%를 차지하였고, 혼인 중인 대상자는 1,871명으로 전체 대상자의 73.75%였고, 별거/이혼/사별에 해당하는 대상자는 658명으로 전체 대상자의 25.94%였다.

주관적 건강상태를 살펴보면 주관적 건강상태가 좋음인 대상자는 452명으로 전체 대상자의 17.82%였고, 주관적 건강상태가 보통인 대상자는 1,273명으로 전체 대상자의 50.18%였고, 주관적 건강상태가 나쁨인 대상자는 812명으로 전체 대상자의 32.01%였다. 음주 여부를 살펴보면 비음주자는 1,413명으로 전체 대상자의 55.70%를 차지하였고, 과거음주자는 367명으로 전체 대상자의

14.47%였고, 현재음주자는 757명으로 전체 대상자의 29.84%였다. 흡연 여부를 살펴보면 비흡연자는 1,833명으로 전체 대상자의 72.25%였고, 과거흡연자는 477명으로 전체 대상자의 18.80%였고, 현재흡연자는 227명으로 전체 대상자의 8.95%였다. 현재 노동 여부를 살펴보면 노동하지 않는 대상자는 1,781명으로 전체 대상자의 70.20%였고, 노동하는 대상자는 756명으로 전체 대상자의 29.80%였다. 가구 총 자산을 살펴보면 저소득에 해당하는 대상자는 1,456명으로 전체 대상자의 56.05%였고, 중산층에 해당하는 대상자는 618명으로 전체 대상자의 21.32%였고, 고소득에 해당하는 대상자는 463명으로 전체 대상자의 22.63%였다. 고혈압 진단 여부를 살펴보면 고혈압 진단받지 않은 대상자는 1,232명으로 전체 대상자의 48.56%였고, 고혈압 진단을 받은 대상자는 1,305명으로 전체 대상자의 51.44%였다. 당뇨 진단 여부를 살펴보면 당뇨 진단받지 않은 대상자는 2,002명으로 전체 대상자의 78.91%였고, 당뇨 진단받은 대상자는 535명으로 21.09%였다. 보청기 사용 여부를 살펴보면 보청기를 사용하지 않는 대상자는 2,491명으로 전체 대상자의 98.19%였고, 보청기를 사용하는 대상자는 46명으로 전체 대상자의 1.81%였다. 틀니 사용 여부를 살펴보면 틀니를 사용하지 않는 대상자는 1,790명으로 전체 대상자의 70.56%였고, 틀니를 사용하는 대상자는 747명으로 전체 대상자의 29.44%였다.

표 3. 연구대상자의 일반적 특성 (5차 조사 시점 기준)

변수명	분류	설명 모델 분석군	예측 모델 분석군
		n (%)	n (%)
성별	1) 남자	1,678 (42.64)	1,033 (40.72)
	2) 여자	2,257 (57.36)	1,504 (59.28)
연령 그룹	1) 65세 이상 75세 미만	2,135 (54.26)	1,656 (65.27)
	2) 75세 이상 85세 미만	1,507 (38.30)	815 (32.12)
	3) 85세 이상	293 (7.45)	66 (2.60)
최종 학력	1) 초등학교 졸업 이하	2,387 (60.66)	1,473 (58.06)
	2) 중학교 졸업	589 (14.97)	401 (15.81)
	3) 고등학교 졸업	676 (17.18)	477 (18.80)
	4) 대학교 이상 졸업	283 (7.19)	186 (7.33)
가구원 수 그룹	1) 1명	770 (19.57)	417 (16.44)
	2) 2명	2,292 (58.25)	1,581 (62.32)
	3) 3명	466 (11.84)	297 (11.71)
	4) 4명 이상	407 (10.34)	242 (9.54)
결혼 여부	1) 미혼	11 (0.28)	8 (0.32)
	2) 혼인 중	1,253 (31.84)	1,871 (25.94)
	3) 별거/이혼/사별	2,671 (67.88)	658 (73.75)
주관적 건강상태	1) 좋음	633 (16.09)	452 (17.82)
	2) 보통	1,843 (46.84)	1,273 (50.18)
	3) 나쁨	1,459 (37.08)	812 (32.01)

표 3. 연구대상자의 일반적 특성 (5차 조사 시점 기준) (계속)

변수명	분류	설명 모델	분석군	예측 모델	분석군
		n (%)	n (%)	n (%)	n (%)
음주 여부	1) 비음주자	2,181 (55.43)		1,413 (55.70)	
	2) 과거음주자		680 (17.28)		367 (14.47)
	3) 현재음주자		1,074 (27.29)		757 (29.84)
흡연 여부	1) 비흡연자		2,759 (70.11)		1,833 (72.25)
	2) 과거흡연자		779 (19.80)		477 (18.80)
	3) 현재흡연자		397 (10.09)		227 (8.95)
현재 노동 여부	1) 노동하지 않음	2,962 (75.27)		1,781 (70.20)	
	2) 노동함		973 (24.73)		756 (29.80)
가구 총 자산	1) 저소득		2,290 (58.20)		1,456 (56.05)
	2) 중산층		773 (19.64)		618 (21.32)
	3) 고소득		872 (22.16)		463 (22.63)
고혈압 진단 여부	1) 진단받지 않음	1,861 (47.29)		1,232 (48.56)	
	2) 고혈압		2,074 (52.71)		1,305 (51.44)
당뇨 진단 여부	1) 진단받지 않음	3,040 (77.26)		2,002 (78.91)	
	2) 당뇨		895 (22.74)		535 (21.09)
보청기 사용 여부	1) 사용하지 않음	3,831 (97.36)		2,491 (98.19)	
	2) 사용함		104 (2.64)		46 (1.81)
틀니 사용 여부	1) 사용하지 않음	2,656 (67.50)		1,790 (70.56)	
	2) 사용함		1,279 (32.50)		747 (29.44)

2. 혼합 효과 로지스틱 회귀모형의 결과

설명 모델 분석군에서 혼합 효과 로지스틱 회귀모형의 통계적 모형 적합도를 확인하기 위해 AIC 값을 확인하였다. 본 연구에서는 공분산 행렬의 다양한 구조 중 분산성분 구조, 복합 대칭성 구조, 비구조적 구조, 1차 자기회귀 구조, Toeplitz 구조의 AIC 값을 표 4와 같이 비교하였다. 표 4에 따르면 분산성분, 비구조적, Toeplitz의 AIC 값이 4289.10로 동일하였다. 복합 대칭성 구조의 AIC값은 4291.12, 1차 자기회귀 구조의 AIC 값은 4291.10으로 약간 더 높았다. 분산성분 구조, 비구조적 구조, Toeplitz 구조 중 선택하기 위해 AICC, BIC 값을 비교하였으나 AICC 4289.21, BIC 4477.43으로 동일하였다 따라서 3가지 구조 중 가장 적은 매개변수를 사용하는 분산성분 구조를 선택하여 분석에 사용하였다.

표 4. 공분산 행렬 구조의 AIC 값 비교 (설명 모델 분석군)

공분산 행렬 구조	분산성분 (VC)	복합대칭성 (CS)	비구조적 (UN)	1차 자기회귀 (AR(1))	Toeplitz (TOEP)
AIC	4289.10	4291.12	4289.10	4291.10	4289.10
AICC	4289.21	4291.24	4289.21	4291.22	4289.21
BIC	4477.43	4485.73	4477.43	4485.71	4477.43

예측 모델 분석군에서 혼합 효과 로지스틱 회귀모형의 통계적 모형 적합도를 확인하기 위해 AIC 값을 확인하였다. 본 연구에서는 공분산 행렬의 다양한 구조 중 분산성분 구조, 복합 대칭성 구조, 비구조적 구조, 1차 자기회귀 구조, Toeplitz 구조의 AIC 값을 표 5와 같이 비교하였다.

표 5에 따르면 분산성분 구조, 비구조적 구조, Toeplitz 구조의 AIC 값이 2565.53으로 동일하였다. 복합 대칭성 구조의 AIC 값은 2567.51, 1차 자기회귀 구조의 AIC 값은 2567.53으로 약간 높았다. 분산성분 구조, 비구조적 구조, Toeplitz 구조 중 선택하기 위해 AICC, BIC 값을 비교하였으나 AICC 2565.57, BIC 2740.69로 동일하였다. 따라서 3가지 구조 중 가장 적은 매개변수를 사용하는 분산성분 구조를 선택하여 분석에 사용하였다.

표 5. 공분산 행렬 구조의 AIC 값 비교 (예측 모델 분석군)

공분산 행렬 구조	분산성분 (VC)	복합대칭성 (CS)	비구조적 (UN)	1차 자기회귀 (AR(1))	Toeplitz (TOEP)
AIC	2565.53	2567.51	2565.53	2567.53	2565.53
AICC	2565.57	2567.67	2565.67	2567.68	2565.67
BIC	2740.69	2748.51	2740.69	2748.53	2740.69

설명 모델 분석군과 예측 모델 분석군 모두 분산성분 구조의 혼합 효과로지스틱 회귀모형을 적용하였고 그 결과를 표 6에 제시하였다. 오즈비의 95% 신뢰구간이 1을 포함할 경우 * 기호를 통해 통계적으로 유의함을 나타내었다 (표 6). 혼합효과 로지스틱 회귀모형 결과 설명 모델 분석군과 예측 모델 분석군 모두에서 성별, 최종 학력, 결혼 여부, 흡연 여부, 가구 총 자산, 고혈압 진단 여부, 당뇨 진단 여부, 틀니 사용 여부 변수는 유의하지 않았다.

연령 그룹을 살펴보면, ‘85세 이상’ 그룹은 ‘65세 이상 75세 미만’ 그룹에 비해 도움필요군일 오즈비는 설명 모델 분석군에서 2.299배(95% 신뢰구간: 1.176 - 4.492) 높았으며 통계적으로 유의하였지만 예측 모델 분석군에서는 1.531배(95% 신뢰구간: 0.687 - 3.410) 높았고 통계적으로 유의하지 않았다. 이는 ‘85세 이상’ 그룹이 설명 모델 분석군 293명에서 사망, 추적 조사 중단 등의 이유로 예측 모델 분석군 66명으로 감소하여 발생된 결과이다. 가구원 수 그룹을 살펴보면, 가구원 수가 ‘2명’ 인 그룹은 가구원 수가 ‘1명’ 인 그룹에 비해 도움필요군일 오즈비는 예측 모델 분석군에서는 2.130배(95% 신뢰구간: 1.069 - 4.241) 높았고 통계적으로 유의하였으나 설명 모델 분석군에서는 유의하지 않았다. 가구원 수가 ‘3명’ 인 그룹은 가구원 수가 ‘1명’ 인 그룹에 비해 도움필요군일 오즈비는 설명 모델 분석군에서 2.254배(95% 신뢰구간: 1.151 - 4.414) 높았고 예측 모델 분석군에서는 2.437배(95% 신뢰구간: 1.065 - 5.580) 높았으며 두 분석군 모두에서 통계적으로 유의하였다. 주관적 건강상태를 살펴보면, ‘나쁨’ 그룹은 ‘좋음’ 그룹에 비해 도움필요군일 오즈비는 설명 모델 분석군에서 4.805배(95% 신뢰구간: 2.628 - 8.785) 높았고 예측 모델 분석군에서는 3.266배(95% 신뢰구간: 1.616 - 6.600) 높았으며 두 분석군 모두에서 통계적으로 유의하였다. 음주 여부를 살펴보면, ‘현재음 주자’ 그룹은 ‘비음주자’ 그룹에 비해 도움필요군일 오즈비는 설명 모델 분석군에서 0.188배(95% 신뢰구간: 0.081 - 0.432) 낮았고 예측 모델 분석군에서는 0.299배(95% 신뢰구간: 0.116 - 0.772) 낮았으며 두 분석군 모두에서 통계적으로 유의하였다. 현재 노동 여부를 살펴보면, ‘노동함’ 그룹은 ‘노동하지 않음’ 그룹에 비해 도움필요군일 오즈비는 설명 모델 분석군에서 0.037배(95% 신뢰구간: 0.009 - 0.147) 낮았고 예측 모델 분석군에서는 0.045배(95% 신뢰구간: 0.011 - 0.195) 낮았으며 두 분석군 모두에서 통계적으로 유의하였다.

보청기 사용 여부를 살펴보면, ‘사용함’ 그룹은 ‘사용하지 않음’ 그룹에 비해 도움필요군일 오즈비는 설명 모델 분석군에서 2.036배(95% 신뢰구간: 1.153 - 3.596) 높았고 예측 모델 분석군에서는 1.983배(95% 신뢰구간: 1.021 - 3.852) 높았으며 두 분석군 모두에서 통계적으로 유의하였다.

표 6. 혼합 효과 로지스틱 회귀모형의 결과

변수	Odds ratio (95% CI)	Odds ratio (95% CI)
성별		
남자	Ref.	Ref.
여자	0.497 (0.229 - 1.078)	0.464 (0.185 - 1.165)
연령 그룹		
65세 이상 75세 미만	Ref.	Ref.
75세 이상 85세 미만	0.753 (0.435 - 1.303)	0.627 (0.332 - 1.183)
85세 이상	2.299 (1.176 - 4.492)*	1.531 (0.687 - 3.410)
최종 학력		
초등학교 졸업 이하	Ref.	Ref.
중학교 졸업	0.610 (0.281 - 1.325)	0.534 (0.212 - 1.342)
고등학교 졸업	0.639 (0.299 - 1.366)	0.489 (0.197 - 1.213)
대학교 이상 졸업	0.563 (0.179 - 1.765)	0.271 (0.057 - 1.278)
가구원 수 그룹		
1명	Ref.	Ref.
2명	1.598 (0.900 - 2.835)	2.130 (1.069 - 4.241)*
3명	2.254 (1.151 - 4.414)*	2.437 (1.065 - 5.580)*
4명 이상	1.738 (0.873 - 3.459)	2.006 (0.877 - 4.591)

* Statistically significant.

표 6. 혼합 효과 로지스틱 회귀모형의 결과 (계속)

변수	Odds ratio (95% CI)	Odds ratio (95% CI)
결혼 여부		
미혼	Ref.	Ref.
혼인 중	0.814 (0.004 - 188.138)	0.364 (0.003 - 45.601)
별거/이혼/사별	1.673 (0.007 - 385.050)	0.769 (0.006 - 95.566)
주관적 건강상태		
좋음	Ref.	Ref.
보통	0.538 (0.281 - 1.031)	0.597 (0.283- 1.260)
나쁨	4.805 (2.628 - 8.785)*	3.266 (1.616 - 6.600)*
음주 여부		
비음주자	Ref.	Ref.
과거음주자	0.918 (0.497 - 1.696)	0.846 (0.405 - 1.770)
현재음주자	0.188 (0.081 - 0.432)*	0.299 (0.116 - 0.772)*
흡연 여부		
비흡연자	Ref.	Ref.
과거흡연자	1.333 (0.658 - 2.703)	1.037 (0.443 - 2.426)
현재흡연자	1.015 (0.342 - 3.014)	0.722 (0.203 - 2.572)
현재 노동 여부		
노동하지 않음	Ref.	Ref.
노동함	0.037 (0.009 - 0.147)*	0.045 (0.011 - 0.195)*

* Statistically significant.

표 6. 혼합 효과 로지스틱 회귀모형의 결과 (계속)

변수	Odds ratio (95% CI)	Odds ratio (95% CI)
가구 총 자산		
저소득	Ref.	Ref.
중산층	0.951 (0.645 - 1.402)	0.760 (0.477 - 1.211)
고소득	1.351 (0.780 - 2.339)	1.106 (0.583 - 2.098)
고혈압 진단 여부		
진단받지 않음	Ref.	Ref.
고혈압	1.021 (0.629 - 1.657)	0.811 (0.461 - 1.428)
당뇨 진단 여부		
진단받지 않음	Ref.	Ref.
당뇨	1.142 (0.695 - 1.875)	1.316 (0.732 - 2.365)
보청기 사용 여부		
사용하지 않음	Ref.	Ref.
사용함	2.036 (1.153 - 3.596)*	1.983 (1.021 - 3.852)*
틀니 사용 여부		
사용하지 않음	Ref.	Ref.
사용함	1.371 (0.986- 1.906)	1.282 (0.869 - 1.892)

* Statistically significant.

3. 구축된 모형의 성능 비교

혼합 효과 로지스틱 회귀모형과 LSTM 모델과 GRU 모델 사이의 성능 비교를 위해 AUC값, 정확도, 오분류율을 지표로 사용하였다. 전체 모형에 대한 성능 지표는 표 9에 제시하였다.

혼합 효과 로지스틱 회귀모형은 AUC 0.6161, 정확도 0.9549, 오분류율 0.0451이었다. LSTM 모델은 AUC 0.7395, 정확도 0.9664, 오분류율 0.0336이었다. GRU 모델은 AUC 0.7516, 정확도 0.9664, 오분류율 0.0336이었다. AUC값은 GRU 모델>LSTM 모델>혼합 효과 로지스틱 회귀모형 순으로 높았다. 정확도는 LSTM 모델=GRU 모델>혼합 효과 로지스틱 회귀모형 순으로 높은 것을 확인하였다. 오분류율은 정확도와 반대로 혼합 효과 로지스틱 회귀모형>LSTM 모델=GRU 모델 순으로 낮은 것을 확인하였다.

표 7. 예측 모델의 성능 비교

Model	AUC	Accuracy	Misclassification rate
Mixed effects	0.6161	0.9549	0.0451
Logistic regression			
LSTM 모델	0.7395	0.9664	0.0336
GRU 모델	0.7516	0.9664	0.0336

그림 5에서는 ROC curve의 시각화를 통해 각 모델의 성능 비교를 한 눈에 보기 쉽게 정리하였다. 혼합 효과 로지스틱 회귀모형, LSTM 모델, GRU 모델 총 3가지 model의 ROC curve를 한 그래프에 시각화시켜 한 눈에 비교할 수 있도록 하였다.

X축은 False Positive Rate로 실제 음성인 경우를 양성으로 잘못 예측한 비율을 나타내며, Y축은 True Positive Rate로 실제 양성인 경우를 양성으로 올바르게 예측한 비율을 나타낸다. 그래프 가운데에 있는 대각선은 AUC 값이 0.5임을 나타내는 기준선이다.

그림 5을 보았을 때 GRU 모델의 AUC 값은 0.7516으로 3가지 model 중 가장 우수한 성능을 보였다. LSTM 모델의 AUC 값은 0.7395로 3가지 model 중 중간 정도의 성능을 보였지만 False Positive Rate가 0.4 부근에서는 GRU 모델보다 약간 우수한 것을 확인할 수 있었다. 또한, False Positive Rate가 0.7 이상인 구간에서는 GRU 모델과 비슷한 성능을 보이는 것을 알 수 있었다. 혼합 효과 다항 로지스틱 회귀모형의 AUC 값이 0.6161로 3가지 model 중 전반적으로 가장 낮은 성능을 보였다.

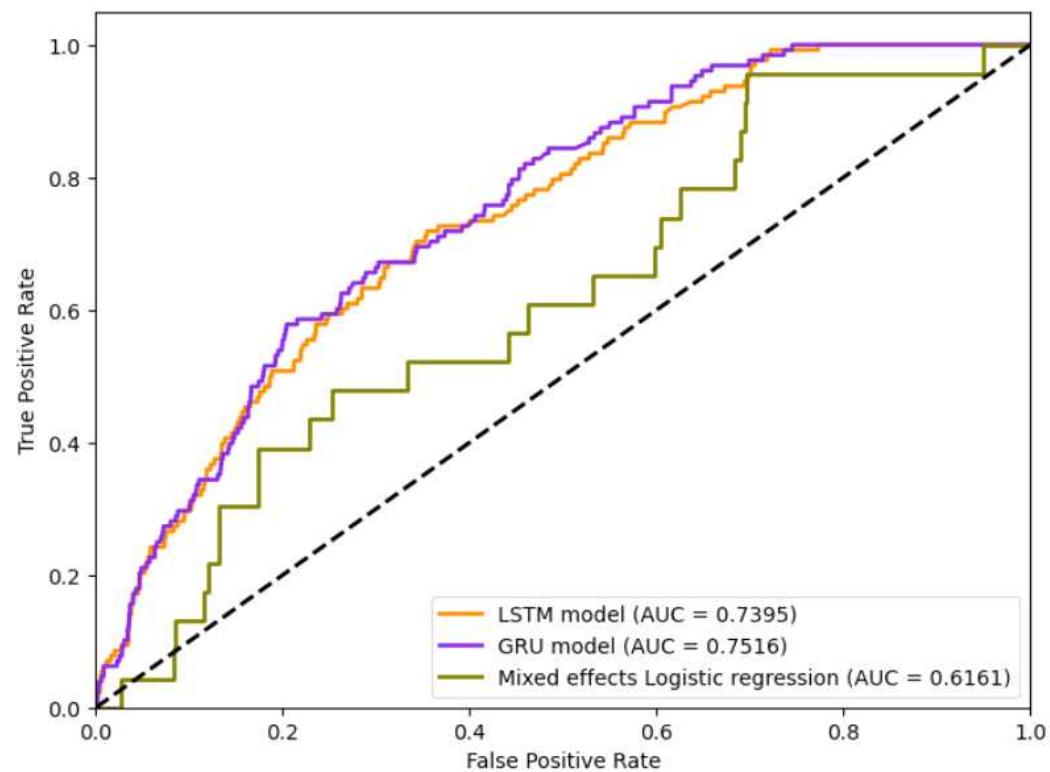


그림 5. ROC Curve

IV . 고찰

1. 연구 결과에 대한 고찰

본 연구는 고령화연구패널(KLoSA) 5~9차 자료를 활용하여 고령인구의 일상 생활수행능력 예측을 위해 혼합효과 로지스틱 회귀모형과 딥러닝 기반 RNN(LSTM, GRU) 모델을 개발하고, 각 모델의 예측 성능을 비교하였다. 분석 결과, 혼합효과 로지스틱 회귀모형의 AUC는 0.6161, LSTM 모델은 0.7395, GRU 모델은 0.7516으로, 딥러닝 기반 RNN 모델이 전통적 통계모형에 비해 우수한 분류 성능을 보였다. 특히 GRU 모델이 가장 높은 AUC 값을 보여, 고령인구의 신체기능 변화 예측에 있어 시계열적 · 비선형적 패턴을 효과적으로 학습할 수 있는 딥러닝 모델의 활용 가능성을 확인하였다.

기존의 고령인구 ADL 예측 연구는 주로 로지스틱 회귀분석, 혼합효과 회귀 모형 등의 방법이 많이 사용되었다 (Oh et al., 2024). 본 연구는 반복측정 자료의 시계열적 특성을 반영한 딥러닝 모델의 적용 가능성을 제시했다는 점에서 기존 연구와 차별성을 가진다. 혼합효과 로지스틱 회귀모형의 예측 성능이 LSTM, GRU 모델에 비해 낮았던 것은, 반복측정 자료에서 시간적 상관성과 비선형적 변화를 충분히 반영하지 못하는 전통적 통계모형의 한계를 시사한다 (George et al., 2015). GRU 모델은 LSTM 모델에 비해 파라미터 수가 적고 학습 속도가 빠르며, 소규모 또는 결측이 많은 데이터에도 강건한 성능을 보인다는 점이 여러 연구에서 보고되고 있다(Chung et al., 2014). 본 연구에서도 결측치 제거 후 분석에서 GRU 모델이 가장 높은 예측력을 보인 것은 이러한 구조적 특성의 영향으로 볼 수 있다.

모델 성능 비교 지표로 사용된 AUC, 정확도, 오분류율은 각각의 장단점이 있으나, AUC는 임계값에 독립적으로 모델의 분류 능력을 평가할 수 있어 불균형 데이터나 다양한 임계값에서의 성능 비교에 적합하다. 본 연구에서도 AUC 기준으로 GRU, LSTM, 혼합효과 로지스틱 회귀모형 순으로 성능이 높게 나타났으며, 이는 의료 및 노인 건강 분야에서 AUC가 모델 성능 평가의 표준 지표로 활용되는 최근 경향과도 부합한다(Anne et al., 2022). 본 연구 결과는 타 분야의 시계열 예측 연구와도 일치하는 경향을 보인다. 예를 들어, Soni et al.(2025)은 시계열 변형 예측에서 LSTM 기반 모델이 기존 통계모형보다 우수한 정확도를 기록했다고 보고했으며, Saha et al.(2023)은 대규모 환자 데이터에서 GRU 모델이 복잡한 시계열 패턴을 효과적으로 학습하여 높은 예측력을 보였다고 밝혔다. 또한, 본 연구에서 혼합효과 로지스틱 회귀모형의 주요 예측요인으로 연령, 주관적 건강상태, 현재 노동 여부, 음주 여부, 보청기 사용 여부 등이 유의하게 나타났으며, 이는 기존 연구에서 확인된 일상생활수행능력 저하의 위험요인과도 일치하는 결과이다(Singh et al., 2024). 특히 주관적 건강상태와 노동 여부가 일상생활수행능력 저하와 밀접하게 연관된다는 점은 고령자의 건강관리 및 사회참여 정책 수립에 중요한 시사점을 제공한다.

2. 연구의 제한점과 의의

2.1. 연구의 제한점

본 연구는 다음과 같은 제한점을 가진다. 첫째, 본 연구에서 개발한 예측모델은 특정 시점(2014~2022년)과 특정 집단(만 65세 이상 완전자립군)에 국한되어 있다. 특히 85세 이상 그룹은 설명 모델 분석군 293명에서 사망, 추적조사 중단 등의 이유로 예측 모델 분석군 66명으로 감소하여, 이 집단의 일상 생활수행능력 변화 패턴을 일반화하기 어렵다는 한계가 있다. 이를 해결하기 위해 고령화패널 외 다른 연구 자료를 이용한 추가 연구가 필요하다. 둘째, 본 연구에서 사용한 LSTM, GRU 모델은 중도 탈락된 대상자들을 제외해야만 분석이 가능하였기 때문에 설명 모델 분석군의 일반적 특성과 차이가 발생하였다. 이로 인해 설명 모델 분석군에서는 유의하였던 85세 이상 그룹이 예측 모델 분석군에서는 유의하지 않은 결과가 나타났다. 셋째, 본 연구는 LSTM, GRU 등 RNN 계열 딥러닝 모델과 혼합효과 로지스틱 회귀모형만을 비교 대상으로 선정하여 모델의 예측 결과의 해석에 제한이 있으며 향후 설명 가능한 다양한 모델 간 연구가 필요하다.



2.2. 연구의 의의

그럼에도 불구하고 본 연구의 의의는 다음과 같다. 첫째, 기존 고령자 일상 생활수행능력 예측 연구가 주로 로지스틱 회귀, 혼합효과 회귀 등의 분석을 사용한 반면, 본 연구는 LSTM, GRU 등 RNN 계열 딥러닝 모델을 적용하여 시계열 데이터의 장기 의존성과 비선형적 변화를 효과적으로 반영하였다. 둘째, 본 연구는 KLoSA 패널의 장기 추적자료를 활용하여 실제 고령인구의 일상생활 수행능력 저하를 예측하는 데 있어 딥러닝 기반 시계열 모델의 적용 가능성을 제시하였다. 이는 고령자 개개인의 건강위험 예측 및 맞춤형 건강관리 정책 수립, 조기 개입 전략 마련에 중요한 근거를 제공할 수 있다. 또한, 본 연구에서 제시한 분석 프레임워크는 향후 인지기능, 낙상 위험 등 타 건강지표 예측 연구에도 적용 가능하며, 반복측정 건강자료의 예측모형 개발에 있어 표준화된 접근법으로 활용될 수 있다.

V . 결론

본 연구는 고령화연구패널(KLoSA) 5~9차 추적자료를 활용하여, 고령인구의 일상생활수행능력(ADL) 예측을 위한 혼합효과 로지스틱 회귀모형과 딥러닝 기반 순환신경망(LSTM, GRU) 모델을 개발하고, 각 모델의 예측 성능을 비교하였다. 연구 결과, 딥러닝 기반 순환신경망 모델(LSTM, GRU)이 전통적 통계모형인 혼합효과 로지스틱 회귀모형에 비해 더 우수한 예측력을 보였으며, 특히 GRU 모델이 가장 높은 AUC(0.7516)와 정확도를 기록하였다. 혼합효과 로지스틱 회귀모형은 반복측정 자료의 시간적 상관성, 비선형적 변화를 충분히 반영하지 못하는 한계로 인해 예측 성능이 상대적으로 낮았다. 반면, LSTM과 GRU 모델은 시계열 데이터의 장기 의존성과 비선형적 패턴을 효과적으로 학습하여, 고령인구의 ADL 저하 예측에 있어 높은 정확도와 신뢰성을 보였다. 특히, GRU 모델의 구조적 단순성과 학습 효율성은 반복측정 건강자료와 같이 표본 크기에 제한이 있거나 결측 문제가 심화된 데이터에 적용할 때 더욱 두드러진 성능 향상을 가져왔다.

본 연구는 반복측정된 고령자 건강자료의 예측에 있어 딥러닝 기반 시계열 모델의 활용 가능성을 실증적으로 제시하였으며, 이는 고령인구의 신체기능 저하 예측 및 맞춤형 건강관리 정책 수립에 중요한 근거를 제공한다. 앞으로는 다양한 건강지표 및 다른 데이터를 활용한 추가 연구, 설명 가능한 기법을 적용한 해석 가능성 강화, 실제 임상 및 정책 현장 적용을 위한 후속 연구가 필요하다. 본 연구는 고령사회 진입에 따른 건강위험 예측 및 예방적 건강관리 정책 수립에 있어, 데이터 기반 맞춤형 접근의 중요성을 강조하며, 딥러닝 기반 시계열 모델의 실용적 가능성과 학문적 의의를 동시에 제시하였다.

참고문헌

Anne A H de Hond EWS, Ben van Calster. Interpreting area under the receiver operating characteristic curve. THE LANCET Digital Health 2022;4(12):853–5.

Ashish Kharel ZZ, Devinder Kaur Long Short-Term Memory (LSTM) Based Deep Learning Models for Predicting Univariate Time Series Data International Journal of Machine Learning 2024;14:30.

Ayushi Das PD. Application of machine learning in measurement of ageing and geriatric diseases: a systematic review. BMC Geriatrics 2023;23:841.

Cavanaugh JE NA. The Akaike information criterion:Background, derivation, properties, application, interpretation, and refinements. WIREs Comput Stat 2019:1460.

Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. Doha, Qatar: Association for Computational Linguistics, 2014.

Cho SJ, Watson D, Jacobs C, Naveiras M. A Markov Mixed-Effect Multinomial Logistic Regression Model for Nominal Repeated Measures with an Application to Syntactic Self-Priming Effects. *Multivariate Behav Res* 2021;56(3):476-95.

George B, Aban I. Selecting a separable parametric spatiotemporal covariance structure for longitudinal imaging data. *Stat Med* 2015;34(1):145-61.

Chung JY, Cho KH, Yoshua Bengio. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *NIPS* 2014.

Li J, Wong WK. Selection of covariance patterns for longitudinal data in semi-parametric models. *Stat Methods Med Res* 2010;19(2):183-96.

Nick TG, Campbell KM. Logistic regression. *Methods Mol Biol* 2007;404:273-301.

Oh E, Moon S, Hong GS. Longitudinal trends and predictors of limitations in activities of daily living in community-dwelling older adults: evidence from the KLoSA study. *Front Public Health* 2024;12:1485732.

Pencina MJ, D'Agostino RB, Sr. Evaluating Discrimination of Risk Prediction Models: The C Statistic. *JAMA* 2015;314(10):1063-4.

Saha A, Samaan M, Peng B, Ning X. A Multi-Layered GRU 모델 for COVID-19 Patient Representation and Phenotyping from Large-Scale EHR Data. *ACM BCB* 2023;2023.

Saito M, Haraguchi H, Nakajima I, Fukuda S, Zhu C, Masuya N, Matsumoto K, Yoshikawa Y, Tanaka T, Kishigami S, Matsumoto L. A machine learning model for predicting fertilization following short-term insemination using embryo images. *Reprod Med Biol* 2025;24(1):e12649.

Shi H, Ding K, Yang XT, Wu TF, Zheng JY, Wang LF, Zhou BY, Sun LP, Zhang YF, Zhao CK, Xu HX. Prediction of BRAF and TERT status in PTCs by machine learning-based ultrasound radiomics methods: A multicenter study. *J Clin Transl Endocrinol* 2025;40:100390.

Singh A, Akkala S, Nayak M, Kotlo A, Poondla N, Raza S, Stankovich J, Antony B. Impact of Pain on Activities of Daily Living in Older Adults: A Cross-Sectional Analysis of Korean Longitudinal Study of Aging (KLoSA). *Geriatrics (Basel)* 2024;9(3).

Soni R, Alam MS, Vishwakarma GK. Prediction of InSAR deformation time-series using improved LSTM deep learning model. *Sci Rep* 2025;15(1):5333.

Ten Have TR, Kunselman AR, Pulkstenis EP, Landis JR. Mixed effects logistic regression models for longitudinal binary response data with informative drop-out. *Biometrics* 1998;54(1):367-83.

Tsao CW, Vasan RS. Cohort Profile: The Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. *Int J Epidemiol* 2015;44(6):1800-13.

Wang LY, Hu ZY, Chen HX, Zhu H, Zhou CF, Zhang RX, Tang ML, Hu XY. Systematic review of longitudinal studies on daily health behavior and activity of daily living among older adults. *Front Public Health* 2025;13:1419279.

Won CW, Rho YG, Kim SY, Cho BR, Lee YS. The validity and reliability of Korean Activities of Daily Living (K-ADL) scale. *Journal of the Korean Geriatrics Society* 2002;6(2):98-106.

Won CW, Rho YG, SunWoo D, Lee YS. The validity and reliability of Korean Instrumental Activities of Daily Living (K-IADL) scale. *Journal of the Korean Geriatrics Society* 2002;6(4):273-80.

Won CW, Yang KY, Rho YG, Kim SY, Lee EJ, Yoon JL, Cho KH, Shin HC, Cho BR, Oh JR. The development of Korean activities of daily living (K-ADL) and Korean instrumental activities of daily living (K-IADL) scale. Journal of the Korean Geriatrics Society 2002;6(2):107-20.

Won CW, Yang KY, Rho YG, Kim SY, Lee EJ, Yoon JL, Cho KH, Shin HC, Cho BR, Oh JR, Yoon DK, Lee HS, Lee YS. The Development of Korean Activities of Daily Living(K-ADL) and Korean Instrumental Activities of Daily Living(K-IADL) Scale. J Korean Geriatr Soc 2002;6(2):107-20.

Yuan DJ, Yang ZX. [Application of nonlinear mixed models in Logistic regression with random effect in clinical trials]. Nan Fang Yi Ke Da Xue Xue Bao 2010;30(8):1923-5, 9.

Zabor EC, Reddy CA, Tendulkar RD, Patil S. Logistic Regression in Clinical Studies. Int J Radiat Oncol Biol Phys 2022;112(2):271-7.

한국고용정보원 고용동향분석팀. 고령화연구패널조사 유저가이드. 2024.

김유진. 한국인 신체기능상태평가(ADL, IADL)예측모형 개발과 미래 예측 적용 연구. 서울: 연세대학교 보건대학원, 2016.

김철립. 디지털 트윈 모델을 적용한 스마트 축사의 인공지능 시뮬레이션

시스템에 관한 연구. 서울: 한성대학교 대학원, 2023.

김호현. LSTM/GRU 순환신경망을 이용한 시계열데이터 예측. 서울:

한국방송통신대학교, 2017.

박수지, 박병선. Testing reliability and measurement invariance of

K-ADL. *Health and Social Welfare Review* 2017;37(4):98-124.

박주성. 포괄적 노인평가와 접근법. *포괄적 노인평가와 접근법* 2005;6(3).

보건복지부, 한국건강증진개발원. 국민건강증진종합계획(HP2030). 2021.

유건령. 순환신경망을 이용한 생존 분석 모형의 성능 비교. 서울: 고려대학교
대학원, 2020.

이소영. 우리나라 노인의 최장기 직종 및 종사상 지위와

일상생활수행능력(ADL) 완전 자립 여부의 관련성. 서울: 연세대학교
보건대학원, 2024.

통계청. 2024 고령자 통계. 2024.

통계청. 2024 장래인구추계 통계정보보고서. 2024.



홍갑표. 지역사회노인들의 일상생활수행능력과 의료이용수준 변화. 서울:
연세대학교 보건대학원, 2003.

ABSTRACT

Development of Predictive Models for Activities of Daily Living (ADL) in Older Adults Using Recurrent Neural Network (RNN) Techniques

This study utilized data from the 5th to 9th waves of the KLoSA to develop and compare both statistical and machine learning-based models for predicting ADL among older adults. While prior research has mainly employed traditional statistical methods—such as logistic regression and mixed-effects models—this study introduces a novel approach by applying deep learning-based RNN model, specifically LSTM and GRU models, to address the long-term dependency challenges of time-series data in forecasting changes in ADL.

The study population comprised individuals aged 65 and older who were fully independent in ADL at the baseline 5th wave in 2014 and were followed through the 9th wave in 2022. Participants who dropped out during any wave from the 5th to the 9th were excluded from the predictive modeling analysis.

Key independent variables included gender, age, education level, household size, marital status, subjective health status, alcohol and smoking habits, employment status, total household assets, diagnoses of hypertension or diabetes, and use of hearing aids or dentures. Predictive models included mixed-effects logistic regression, LSTM, and GRU architectures, evaluated using AUC, accuracy, and misclassification rate, and results showed significant associations between ADL decline and age, subjective health, employment, alcohol use, and hearing aid use, with the GRU model achieving the highest AUC (0.7516) and both deep learning models outperforming the statistical approach (AUC: 0.6161).

This study demonstrates the superiority of deep learning-based time-series models for predicting longitudinal health outcomes in older adults, providing a foundation for forecasting functional decline and informing personalized health management strategies, and suggests further research on ADL recovery prediction and interpretable model development.

Key words : ADL, mixed-effects logistic regression, RNN model, LSTM model, GRU model