

Original Paper

# Explainable Machine Learning for Assessing Digital Health Literacy in Older Adults: Validation and Development of a Two-Stage Model Integrating Performance-Based and Self-Assessed Indicators

Choonghee Park<sup>1</sup>, MA; Jiyeon Park<sup>2</sup>, MA; Seora Kim<sup>1</sup>, MA; Ye Seul Bae<sup>3,4</sup>, MD, PhD; Jae-Heon Kang<sup>4</sup>, MD, PhD; Tae-Min Kim<sup>1,2</sup>, MD, PhD; Ji-Won Chun<sup>1,2</sup>, PhD

<sup>1</sup>Department of Medical Informatics, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

<sup>2</sup>The Catholic Medical Center Institute for Basic Medical Science, The Catholic University of Korea Catholic Medical Center, Seoul, Republic of Korea

<sup>3</sup>Big Data Research Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

<sup>4</sup>Department of Family Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

**Corresponding Author:**

Ji-Won Chun, PhD

Department of Medical Informatics

College of Medicine

The Catholic University of Korea

222 Banpo-daero, Seocho-gu

Seoul, 06591

Republic of Korea

Phone: 82 3147 9727

Email: [jwchun@catholic.ac.kr](mailto:jwchun@catholic.ac.kr)

## Abstract

**Background:** Digital health literacy (DHL) is the ability to locate, understand, evaluate, and apply health information in digital environments. It is essential for older adults to effectively engage with contemporary health care. However, existing DHL assessments primarily rely on self-reported measures, which are susceptible to subjective bias and often fail to capture actual performance. There is a need for a comprehensive, data-driven approach that integrates objective performance indicators with self-assessments to accurately predict and explain DHL levels in older adults.

**Objective:** This study develops and validates a machine learning approach to predict DHL levels in older adults by integrating performance-based and self-assessed evaluations.

**Methods:** We applied a 2-stage methodological framework using 2 independent datasets. In the first stage, to identify performance-based determinants, we assessed actual digital and information comprehension in a separate pilot cohort of 30 older adults (aged 60-74 years). In parallel, to measure self-reported DHL, we conducted an online survey with a distinct group of 1000 older adults (aged 55-74 years) using the Digital Health Literacy Scale and the Korean version of the eHealth Literacy Scale (KeHEALS). Bayesian linear regression was applied to both datasets to identify significant explanatory variables. In the second phase, we trained and validated a binary classification model to predict KeHEALS levels using the survey dataset (n=1000), leveraging the features identified in the first stage. Five machine learning algorithms were evaluated, and the best-performing model was interpreted using Shapley Additive Explanations (SHAP) analysis.

**Results:** In the pilot performance-based assessment, using a greater number of electronic devices and having higher educational attainment were positively associated with comprehension, whereas alcohol intake showed a negative association. In the self-assessed survey data, key correlates included interest in health-related apps, self-care confidence, age, smoking, alcohol intake, number of devices used, and exercise frequency. Among the machine learning models, categorical boosting demonstrated the most balanced performance (accuracy 0.785, precision 0.769, F1-score 0.765, area under the receiver operating characteristic curve 0.835), outperforming the dummy classifier (accuracy 0.540). SHAP analysis indicated that self-care confidence, health information search, interest in health-related apps, number of electronic devices used, and exercise frequency were the strongest positive contributors to high-DHL predictions, whereas older age and lifestyle factors (alcohol intake, smoking) contributed negatively.

**Conclusions:** By explicitly integrating performance-based and self-assessed indicators within an explainable machine learning framework, this study demonstrates that DHL in older adults is influenced by both digital engagement and health management

factors. These findings suggest that the proposed framework can serve as a structured approach for evaluating DHL in older adults and inform the design of personalized digital health interventions in clinical and community settings.

(*JMIR Med Inform 2026;14:e86171*) doi: [10.2196/86171](https://doi.org/10.2196/86171)

## KEYWORDS

digital health literacy; eHealth literacy; digital health care; mHealth; machine learning

## Introduction

### Background

The rapid advancement of digital technologies and smart devices has transformed everyday life, particularly in the health care domain. Digital health technologies such as electronic health records, telemedicine, and mobile health (mHealth) apps are increasingly recognized as essential tools for enhancing access to health information, supporting self-management, and improving the efficiency and quality of health care delivery [1]. These technologies empower individuals to better manage their health and facilitate personalized, patient-centered care [2]. Furthermore, community-driven digital health systems have demonstrated their potential to establish sustainable and equitable models of care [3].

However, to fully benefit from digital health technologies, individuals must have a sufficient level of digital health literacy (DHL). DHL refers to the ability to locate, understand, evaluate, and apply health-related information in digital environments to make informed decisions and manage health [4]. Higher DHL levels have been associated with improved health outcomes, including better chronic disease management, engagement in preventive behaviors, increased use of health care services, enhanced quality of life, and reduced risk of depression [5-8]. Furthermore, DHL plays a critical role in narrowing the digital divide and promoting equitable access to health resources across diverse populations.

According to the 2024 Digital Information Divide Survey in Korea, the digital informatization level among older adults corresponded to 71.4% of that observed in the total population. While digital access has largely equalized, reaching 95.3% of the general population's level, their digital competencies and utilization indices were relatively lower (55.9% and 75.0%, respectively) [9]. These scores are more likely the result of cohort effects and differences in the availability of educational opportunities, health status, and social and technical support, rather than problems caused by aging [10,11]. Consequently, interventions aimed at enhancing DHL in older populations should be evidence-based and sensitive to the diversity of digital experiences in later life, focusing on modifiable barriers related to access, education, and available support rather than chronological age alone.

The majority of existing studies have relied on self-assessment surveys to measure DHL [12-17]. However, such instruments often fail to capture psychosocial dimensions, such as social support, and may not accurately reflect individuals' actual digital competencies [18]. This underscores the need for a more comprehensive and empirically grounded approach to DHL assessment. Machine learning offers a robust framework for

integrating diverse variables, including sociodemographic characteristics, health behaviors, and psychosocial factors, into multidimensional predictive models capable of capturing complex nonlinear interactions [19].

In this study, we employed a Bayesian estimation method to identify key factors associated with DHL-related outcomes, specifically leveraging a wide range of input features, including psychosocial attributes. Subsequently, we developed machine learning models to predict DHL levels and applied Shapley Additive Explanations (SHAP), an explainable artificial intelligence technique, to enhance the interpretability and transparency of the model's outputs [20]. This 2-stage approach aims to provide a more accurate and comprehensive understanding of DHL in older adults and offers practical insights into the design of personalized digital health education and intervention strategies.

### Assessment Tool for Measuring Digital Health Literacy: eHealth Literacy Scale

Norman and Skinner's [21] foundational work conceptualized eHealth literacy as a composite of 6 core literacies: traditional literacy, information literacy, media literacy, health literacy, scientific literacy, and computer literacy. This framework is grounded in the "Lily Model," which categorizes skills into analytical and context-specific domains. As digital health technologies have advanced, the original construct of eHealth literacy has broadened and is now more commonly used within the wider notion of DHL.

At the same time, several conceptual and methodological limitations of the eHealth Literacy Scale (eHEALS) have been identified and should be considered when the instrument is used. As a self-report measure, eHEALS assesses subjective confidence rather than objectively measured performance, so scores may diverge from actual digital and health-related skills [22,23]. Factor-analytic studies have also reported inconsistent dimensional structures across populations and languages, indicating that the underlying constructs may vary by cultural and demographic context and that direct score comparisons or universal cutoffs should be interpreted with caution [24]. Nevertheless, eHEALS has been translated and validated in numerous countries, demonstrates consistently high internal consistency, and remains a pragmatic core instrument for assessing perceived eHealth literacy in large-scale surveys and clinical research [24-33].

Using eHEALS, Berkowsky [17] assessed the DHL levels of adults aged 65 and older in California and examined the impact of sociodemographic characteristics and digital usage patterns. Regression analyses indicated that higher educational attainment, more frequent internet use, and engagement in a wider range of online activities (eg, email, online shopping, social media)

were associated with higher DHL levels, whereas basic demographic variables such as age, gender, and income were not significant predictors. These findings suggest that actual digital engagement may be more critical than demographic background in shaping perceived DHL.

Similarly, Kyaw et al [34] investigated DHL among 434 adults aged 65 and older in urban settings, focusing on the influence of social and digital factors. Their findings revealed that older age was negatively associated with DHL, whereas higher income, extended internet use, and active participation in social media showed positive associations. Notably, older adults who used social media exhibited approximately 4-fold higher DHL levels than nonusers, and those using the internet for more than 3 hours per day demonstrated up to a 35-fold increase in DHL levels. These findings suggest that improving DHL among older adults requires not only the provision of digital technologies but also strategies to promote real-world digital engagement and enhance social participation in digital environments.

### **Assessment Tool for Measuring Digital Health Literacy: Digital Health Literacy Scale**

Early definitions of eHealth literacy focused primarily on individuals' abilities to search for and understand health information on the internet. However, today's digital environment increasingly demands interactive engagement through platforms such as social media, mHealth apps, and telemedicine. Consequently, the concept has evolved into DHL, which encompasses not only information seeking but also the skills to evaluate, apply, and interact with digital health content across diverse platforms [35].

In line with this expanded conceptualization, the Digital Health Literacy Scale (DHLS) was developed as a multidimensional tool to comprehensively measure DHL. The DHLS evaluates competencies across 4 domains: use of digital devices, understanding of health information, use and decision of health information, and use intention [36]. The use of digital devices domain assesses individuals' proficiency in operating digital tools and accessing health-related services. The understanding of health information domain measures the ability to comprehend online content, such as medication instructions and test results. The use and decision of health information domain evaluates both the application of health information in decision-making and the critical judgment of its credibility. The use intention domain captures users' motivation and willingness to engage with digital health platforms and services. It was specifically designed to reflect the behavioral, cognitive, and motivational characteristics commonly observed in older adults, particularly those vulnerable to the digital divide. By independently assessing both digital and health literacy and

integrating them into a unified DHL construct, the DHLS facilitates a comprehensive and multidimensional assessment of digital health literacy. This approach helps identify differences in digital health literacy across various skill domains, from basic digital operations to advanced critical thinking, providing actionable insights into the specific competencies that require support among older adults.

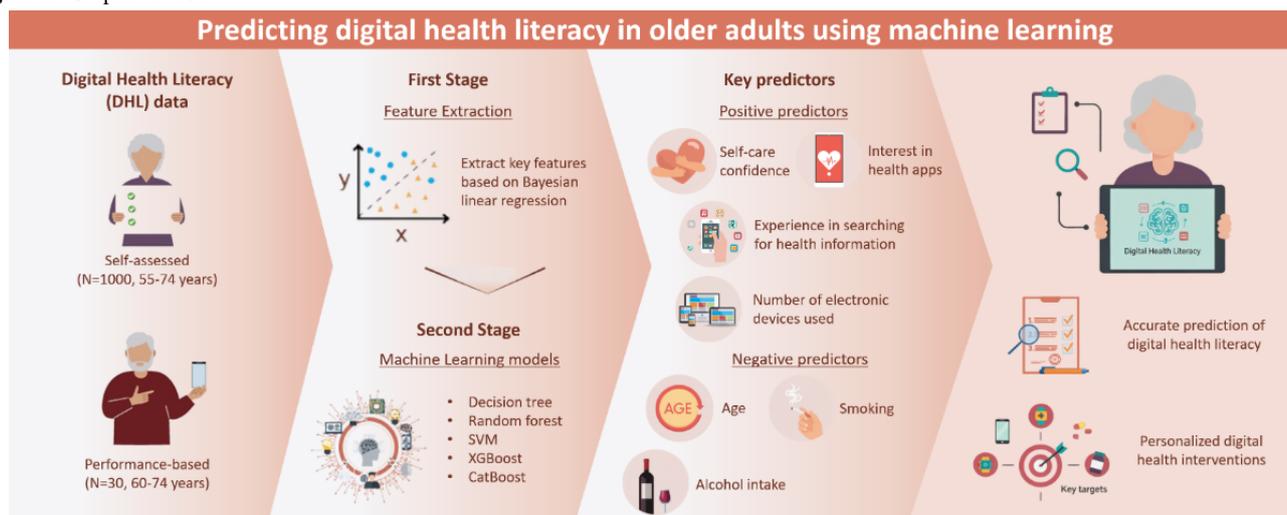
### **Machine Learning Approach for Assessing Digital Health Literacy**

Given what we discussed so far, eHEALS and DHLS are widely used self-report measures for evaluating digital health literacy. Most studies employing these tools use univariate analyses to examine the effects of individual variables independently [37-39]. However, because self-assessment instruments are inherently grounded in subjective perceptions, they have substantial limitations in capturing actual performance and real-world information behaviors. This concern aligns with extensive evidence from cognitive and behavioral research demonstrating that self-assessments are often systematically biased and display only modest associations with objective indicators of competence [40]. Moreover, they fail to account for interactions among variables or evaluate the relative importance of each factor [41,42]. These limitations emphasize the need to develop data-driven approaches that use diverse individual-level characteristics to accurately predict DHL.

To address these limitations, this study adopted a comprehensive 2-stage machine learning framework that integrates performance-based and self-assessed indicators of DHL. Machine learning is a statistical approach that predicts and classifies individual conditions by learning complex, nonlinear patterns from multidimensional data [43]. Research suggests that machine learning algorithms, including Extreme Gradient Boosting (XGBoost), random forest, and logistic regression, can identify key sociodemographic factors influencing general health literacy [44]. Existing studies have demonstrated the effectiveness of machine learning in uncovering intricate variable interactions and enhancing predictive accuracy [45-47], reflecting a broader shift toward objective data-driven assessment in health literacy research.

This study integrated performance-based and self-reported assessments to identify key features associated with DHL. Based on these significant features, a prediction model was developed using a machine learning algorithm. Furthermore, we used SHAP to interpret the contribution of each feature, thereby proposing an explainable artificial intelligence approach for medical apps. This approach aims to effectively predict DHL in older adults and provide a foundation for developing personalized intervention strategies (Figure 1).

Figure 1. Graphical abstract.



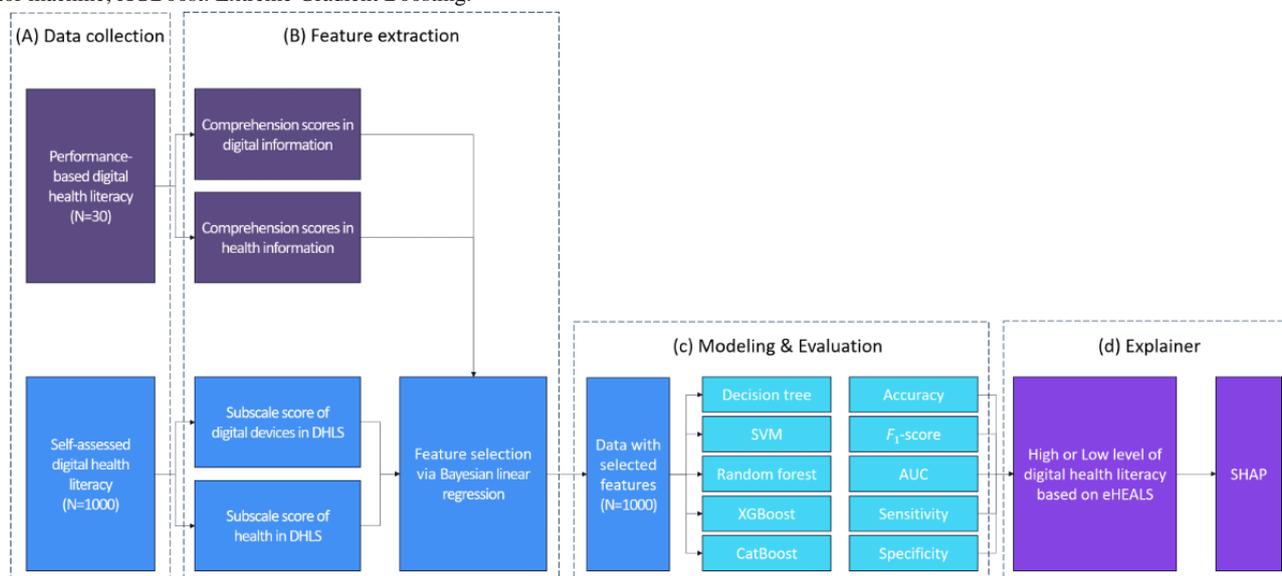
## Methods

### Overview

In this study, we designed a 2-stage machine learning framework integrating self-assessed (n=1000) and performance-based (n=30) evaluations (Figure 2). In the initial phase, using a performance-based task, digital literacy and health care information scores were obtained (Figure 2A). Subsequently, Bayesian linear regression analysis was performed to identify explanatory variables significantly associated with DHL (Figure 2B). The performance-based evaluation was employed to

enhance the explanatory power of the variables and was excluded from the training phase of the DHL predictive model. In the second phase, using a dataset of 1000 self-assessments and the characteristics identified in the initial phase, we constructed prediction models to classify DHL levels (high vs low; Figure 2C). We conducted comparative evaluations of multiple machine learning models, including decision tree, support vector machine (SVM), random forest, XGBoost, and categorical boosting (CatBoost). In addition, SHAP analysis was used to quantitatively assess the impact of significant variables on DHL predictions, thereby enhancing the interpretation of the model’s predicted results (Figure 2D).

Figure 2. Overview of the framework used to predict digital health literacy level. AUC: area under the receiver operating characteristic; CatBoost: categorical boosting; DHLs: Digital Health Literacy Scale; eHEALS: eHealth Literacy Scale; SHAP: Shapley Additive Explanations; SVM: support vector machine; XGBoost: Extreme Gradient Boosting.



## Performance-Based Digital Health Literacy

### Participants

To evaluate performance-based DHL, we recruited a total of 30 participants aged 60-74 (mean age 69.1 years). Participants were enrolled using a nonprobability snowball sampling strategy facilitated through local community networks. To ensure data

validity and minimize potential bias, intermediaries who arranged the initial contact were excluded from participation. Trained research staff (CP and SK) conducted a brief telephone screening to confirm eligibility and then scheduled a home visit. During the visits, researchers provided comprehensive information regarding the study’s purpose, procedures, anonymity, data usage, and participants’ rights. Written

informed consent was obtained before administration of the behavioral tasks. Participation was voluntary, and all participants were informed that they could withdraw from the study at any time without penalty.

To examine differences in performance-based DHL, participants were divided into 2 groups using the eHEALS cutoff score as the criterion: low DHL (n=15) and high DHL (n=15), based on

their Korean version of the eHealth Literacy Scale (KeHEALS) scores, as shown in Table 1 [48]. The 2 groups showed no significant differences in age (mean age: low DHL=68.7 years, high DHL=69.4 years;  $P=.19$ ) or gender (male/female: low DHL=2/13, high DHL=5/10;  $P=.34$ ). In addition, no significant difference was found between the groups in years of education ( $P=.49$ ).

**Table 1.** Comparison of performance-based sociodemographic variables by DHL<sup>a</sup> level based on the KeHEALS<sup>b</sup> score.

Characteristics	Groups n (%)		Total n (%)
	Low	High	
<b>Gender</b>			
Male	2 (29)	5 (71)	7 (100)
Female	13 (57)	10 (43)	23 (100)
<b>Age (years)</b>			
60-64	1 (33)	2 (67)	3 (100)
65-69	9 (64)	5 (36)	14 (100)
70-74	5 (38)	8 (62)	13 (100)
<b>Years of education</b>			
Elementary school	3 (60)	2 (40)	5 (100)
Middle school	3 (75)	1 (25)	4 (100)
High school	8 (57)	6 (43)	14 (100)
Community college	0 (0.0)	1 (100)	1 (100)
University or higher	1 (17)	5 (83)	6 (100)

<sup>a</sup>DHL: digital health literacy.

<sup>b</sup>KeHEALS: Korean version of the eHealth Literacy Scale.

### Measurements

Participants engaged in content reading focused on digital literacy and health care information, followed by an identification task designed to assess comprehension. Each reading selection comprised 2 or 3 paragraphs and was displayed on a Samsung Galaxy A15 smartphone equipped with a 163.9-mm display (1080 × 2340 pixels). The smartphone was used as the experimental tool to deliver all visual materials for task performance in a controlled viewing environment. During the identification task, participants were asked to indicate whether 10 digital literacy sentences and 10 health care information sentences derived from the readings were true or false. The average time required to complete the assessment did not exceed 30 minutes. To identify key explanatory variables, scores from this identification task were subsequently used in the initial stage of the DHL framework. Before the assessment, demographic information (gender, age, and years of education), psychosocial factors (living arrangements), digital-related factors (interest in health-related apps, health information search, and number of digital devices used), and health-related factors (alcohol intake, smoking, and exercise frequency) were collected from each participant. Interest in health-related apps was measured using a 5-point Likert scale (1=strongly disagree to 5=strongly agree), capturing the degree of interest in using health-related apps in digital environments. Health information

search was assessed using a binary (yes/no) item asking whether participants had personally searched for health-related information.

To evaluate participants' overall level of digital health literacy, the KeHEALS was administered. KeHEALS is an 8-item self-report instrument designed to measure individuals' perceived ability to locate, evaluate, and apply electronic health information, demonstrating high internal consistency (Cronbach  $\alpha=.89$ ) and good test-retest reliability over a 2-week interval (intraclass correlation 0.80). Each item was rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree), yielding a total score between 8 and 40, with higher scores indicating greater perceived digital health literacy. Participants were classified into 2 groups using a sample-specific threshold based on the median total eHEALS score of 29 in the entire sample: those with a total score greater than 29 were categorized as having a high level of digital health literacy, whereas those with a score of 29 or below were categorized as having a low level. This classification facilitated subsequent analyses by enabling group comparisons according to participants' perceived competencies in engaging with digital health information.

## Self-Assessed Digital Health Literacy

### Participants

To evaluate self-assessed DHL for the predictive modeling phase, survey data were collected from 1000 adults aged 55-74 (mean age 63.6 years). All participants were recruited through an online survey platform operated by a survey agency. Recruitment was facilitated via email and website postings. To ensure a representative sample of the general population, stratified quota sampling was employed based on gender and age groups. Each respondent was assigned a unique access link to ensure single participation and eliminate the possibility of repeated submissions. Given the prospective and anonymous nature of this large-scale survey, the requirement for written informed consent was waived by the Institutional Review Board. Participants were first presented with information on the initial survey screen that provided comprehensive details about the study objectives and procedures. Proceeding to the questionnaire was considered to indicate informed willingness to participate.

To examine group-level differences in DHL, participants were classified into 2 groups based on KeHEALS scores: low DHL (n=563) and high DHL (n=437; Table 2). The high-DHL group had a greater proportion of participants in the younger age range of 55-59 (148/437, 33.9%), whereas the low-DHL group included more participants in the 65-69 (242/563, 43.0%) and 70-74 (80/563, 14.2%) age ranges. Age distribution also differed between the groups (mean age: low DHL=64.2 years, high DHL=62.8 years). Differences were likewise observed in gender distribution (male/female: low DHL=257/306, high DHL=243/194). The proportion of males was higher in the high-DHL group (243/437, 55.6%) than in the low-DHL group (257/563, 45.6%), whereas females constituted the majority in the low-DHL group (306/563, 54.4%). Regarding years of education, a marked difference was found between the 2 groups. In the high-DHL group, more than half of the participants (235/437, 53.8%) had obtained a university degree or higher, whereas the low-DHL group had a greater proportion of high school graduates (271/563, 48.1%) and individuals with less than a middle school education.

**Table 2.** Comparison of self-assessed sociodemographic variables by DHL<sup>a</sup> level based on the KeHEALS<sup>b</sup> score.

Characteristics	Groups, n (%)		Total, n (%)
	Low	High	
<b>Gender</b>			
Male	257 (51.4)	243 (48.6)	500 (100)
Female	306 (61.2)	194 (38.8)	500 (100)
<b>Age (years)</b>			
55-59	134 (47.5)	148 (52.5)	282 (100)
60-64	107 (57.5)	79 (42.5)	186 (100)
65-69	242 (57.9)	176 (42.1)	418 (100)
70-74	80 (70.2)	34 (29.8)	114 (100)
<b>Education</b>			
Elementary school	5 (83.3)	1 (16.7)	6 (100)
Middle school	45 (93.8)	3 (6.3)	48 (100)
High school	271 (66.4)	137 (33.6)	408 (100)
Community college	62 (50.4)	61 (49.6)	123 (100)
University or higher	180 (43.4)	235 (56.6)	415 (100)

<sup>a</sup>DHL: digital health literacy.

<sup>b</sup>KeHEALS: Korean version of the eHealth Literacy Scale.

### Measurements

DHL levels were assessed using the DHLS [36] and KeHEALS [48]. The DHLS consists of 4 subscales: use of digital devices (10 items), understanding of health information (5 items), use and decision of health information (5 items), and use intention (5 items). Participants completed the DHLS questionnaire using a 5-point Likert-type response scale (1=strongly disagree, 5=strongly agree). To determine key explanatory variables related to digital literacy and health care information in older adults, we focused on 3 subscales: use of digital devices, understanding of health information, and use and decision of

health information. Based on a sample-specific threshold corresponding to the median total eHEALS score of 29, eHEALS was dichotomized into 2 groups and used as a predictive variable in the machine learning model. Furthermore, we collected demographic information (gender, age, and years of education), psychosocial factors (living arrangements and social support), digital-related factors (interest in health-related apps, health information search, and number of digital devices used), and health-related factors (alcohol intake, smoking, exercise frequency, and self-care confidence). Self-care confidence was assessed as an indicator of health empowerment in older adults using a single item derived from the Korean

Health Empowerment Scale, which measures confidence in making health management decisions [49]. This item was rated on a 5-point Likert scale (1=strongly disagree to 5=strongly agree).

## Statistical Analysis

### Feature Extraction Based on Bayesian

To identify the most influential factors associated with DHL, we applied Bayesian linear regression to both performance- and survey-based datasets. This approach improves modeling efficiency while enhancing interpretability by incorporating prior-based regularization and probabilistic estimation of regression coefficients, thereby allowing direct assessment of variable importance and uncertainty [50]. Each model was specified separately for each observation unit as follows:

$$Y_i = \beta_0 + \sum_{j=1}^{p'} \beta_j X_{ij} + \varepsilon_i$$

where  $Y_i$  is the standardized score of the  $i$ th participant;  $X_{ij}$  is the value of the  $j$ th feature for the  $i$ th participant;  $\beta_0$  is the intercept;  $\beta_j$  is the regression coefficient for each feature;  $\varepsilon_i$  is the error term; and  $p'$  is the total number of features considered.

The likelihood for the observed data was specified as follows:

$$Y_i | \beta_0, \beta, \sigma^2 \sim N(\mu_i, \sigma^2), \text{ where } \mu_i = \beta_0 + \sum_{j=1}^{p'} \beta_j X_{ij}$$

We assigned weakly informative priors to the model parameters to ensure that posterior inferences were primarily driven by the observed data. Specifically, the intercept term  $\beta_0$  was assigned a prior distribution  $N(0, 5)^2$ , and each regression coefficient  $\beta_j$  received an independent prior  $N(0, 1^2)$  for  $j=1, \dots, p'$ , reflecting an implicit L2 regularization effect. The SD of the error term  $\sigma$  was assigned a HalfCauchy(2) prior.

Posterior distributions were approximated using Markov chain Monte Carlo sampling. For each model, 4 parallel chains were run, each with 2000 tuning steps followed by 2000 sampling steps. Convergence was assessed by monitoring divergence diagnostics throughout the sampling process.

Feature selection was conducted using a Bayesian inference framework. The dataset was first split into training and test sets, and all feature selection procedures were performed exclusively on the training set. In the performance-based dataset, we separately identified predictors of comprehension scores for digital information and health information. In the self-assessed dataset, feature selection was performed for each DHL subscale, namely, use of digital devices and health information understanding and utilization. To prevent data leakage during the Markov chain Monte Carlo sampling procedure used for feature selection, all modeling steps were conducted exclusively on the training dataset. Variables were retained only if the 95% highest density interval (HDI) of their posterior regression coefficients  $\beta$  did not include 0, indicating a statistically significant positive or negative association with the outcome

variable. These statistically validated features were then used exclusively in subsequent stages to construct the final predictive model for DHL.

### Modeling

In this study, final predictive modeling was performed using the self-assessed DHL dataset comprising 1000 participants. The goal of the model was to accurately classify participants into 2 distinct groups: high DHL and low DHL. Before developing the final predictive model, a feature selection process was conducted. This process aimed to identify features strongly associated with both performance-based and self-assessed DHL measures. The resulting feature set was designed to comprehensively represent the various aspects of DHL.

Using this selected feature set, 5 machine learning models—decision tree, random forest, XGBoost, CatBoost, and SVM—were implemented to predict DHL levels [51-55]. During data preprocessing, we applied a structured workflow to handle distinct variable types and used mixed encoding strategies to reflect the measurement properties of each predictor. Nominal categorical variables were encoded using OneHotEncoder. Ordinal variables were generally processed with OrdinalEncoder to preserve their inherent ranking. However, for selected lifestyle variables such as alcohol intake, we applied OneHotEncoder to capture category-specific associations across frequency levels that may not be adequately represented by a single ordered scale. These transformations were integrated into a unified preprocessing pipeline and applied consistently across all data splits, ensuring uniform feature representation and minimizing the risk of data leakage.

For model training and testing, the preprocessed dataset was split into a training set ( $n=800$ ) and a test set ( $n=200$ ). The hyperparameters of each model were optimized using the Optuna framework [56], with 10-fold cross-validation applied to the training set to maximize accuracy. Following hyperparameter optimization, to assess generalization performance, each final model was trained on the entire training set and evaluated on the test set.

### Performance Evaluation

Model performance in predicting DHL levels was quantitatively evaluated using standard metrics for binary classification, including accuracy, precision, sensitivity, specificity,  $F_1$ -score, and area under the receiver operating characteristic curve (ROC AUC). Accuracy indicates the overall proportion of correctly classified cases. Precision reflects the proportion of true high-DHL cases among those predicted to have high DHL. The  $F_1$ -score, which is the harmonic mean of precision and recall, provides a balanced metric that is particularly useful under class imbalance. The ROC AUC assesses the model's ability to distinguish between high- and low-DHL cases across varying decision thresholds. Collectively, these metrics offer a comprehensive evaluation of model performance.

Furthermore, to address the “black box” nature of machine learning and ensure clinical transparency, we used the SHAP framework. SHAP is a model-agnostic method based on cooperative game theory that assigns each feature an importance value for a particular prediction. It calculates the Shapley value,

which represents the average marginal contribution of a feature across all possible coalitions of features [20]. Unlike traditional feature-importance metrics that provide only a global ranking, SHAP offers both global and local interpretability. It allows us to quantify not only the magnitude but also the direction of each feature's impact on the model's output.

In the context of our model for predicting digital health literacy, SHAP provides a granular explanation of how specific demographic, interest in digital health care, health management, and psychosocial features influence the classification of an individual into the high-DHL or low-DHL group. This methodology enables visualization of individual and global feature effects, allowing identification of key determinants of DHL and verification of whether the model's learned patterns are consistent with established clinical research on digital health.

### Ethical Consideration

This study was reviewed and approved by the Institutional Review Board of the Catholic University of Korea, College of Medicine (survey study approval number MIRB 20230825-007; behavioral task study approval number MIRB 20241014-003), and was conducted in accordance with institutional and national guidelines for research involving human participants. For the web-based survey, an online written information sheet described the study purpose, procedures, minimal-risk nature, anonymity, data use, and participants' rights. Data were collected only from individuals who voluntarily agreed to participate after reading this information, with explicit notice that they could withdraw from the survey at any time without penalty. For the behavioral task study, participants received a written information sheet and provided written informed consent before participation. In the

survey study, no personally identifiable information was collected. In the behavioral task study, personal information was collected solely for compensation purposes, stored separately from research data, and destroyed immediately after compensation. In both studies, no personal identifiers were used for analysis or reporting; all analytic datasets were anonymized, encrypted, and stored on secure institutional servers accessible only to the research team. Participants who completed the web-based survey received 500 points (approximately US \$0.35) redeemable within the online platform, and participants in the behavioral task study received KRW 30,000 in cash (approximately US \$20) as compensation for their time. No identifiable images of participants were included in the manuscript or its multimedia appendix.

## Results

### Post Hoc Age-Group Comparison of Selected Features

To verify the validity of applying selected features from a small cohort of individuals aged 60 years and older to the 55-59 age group included in the final model during the feature-extraction phase, we divided participants into 2 age groups and performed statistical comparisons of the selected features. The comparison revealed significant demographic and behavioral differences in several variables (Table 3). Modest but statistically significant age-related differences were observed for interest in health-related apps ( $P=.03$ ), health information search ( $P=.02$ ), smoking ( $P=.004$ ), and alcohol intake ( $P=.007$ ). However, no significant differences were observed in the number of electronic devices used ( $P=.41$ ), self-care confidence ( $P=.19$ ), or exercise frequency ( $P=.86$ ).

**Table 3.** Comparison of self-assessed digital and health-related factors by age groups.

Comparison	Age groups (years), n (%)		Chi-square ( <i>df</i> )	Cramer V	P value
	55-59 (n=282)	60-74 (n=718)			
<b>Interest in health-related apps</b>			10.388 (4)	0.102	.03
Strongly disagree	14 (5.0)	76 (10.6)			
Disagree	46 (16.3)	140 (19.5)			
Neither agree nor disagree	75 (26.6)	169 (23.5)			
Agree	86 (30.5)	196 (27.3)			
Strongly agree	61 (21.6)	137 (19.1)			
<b>Health info search</b>			5.617 (1)	0.075	.02
Yes	224 (79.4)	518 (72.1)			
No	58 (20.6)	200 (27.9)			
<b>Number of electronic devices used</b>			2.902 (3)	0.054	.41
1	100 (35.5)	278 (38.7)			
2	111 (39.4)	291 (40.5)			
3	58 (20.6)	116 (16.2)			
4	13 (4.6)	33 (4.6)			
<b>Self-care confidence</b>			6.095 (4)	0.078	.19
Strongly disagree	0 (0.0)	5 (0.7)			
Disagree	23 (8.2)	58 (8.1)			
Neither agree nor disagree	101 (35.8)	273 (38.0)			
Agree	151 (53.5)	347 (48.3)			
Strongly agree	7 (2.5)	35 (4.9)			
<b>Exercise frequency per week</b>			1.314 (4)	0.036	.859
0	38 (13.5)	81 (11.3)			
1	104 (36.9)	259 (36.1)			
2	63 (22.3)	165 (23.0)			
3	49 (17.4)	133 (18.5)			
4	28 (9.9)	80 (11.1)			
<b>Smoking</b>			13.192 (3)	0.115	.004
Never smoke	158 (56.0)	371 (51.7)			
Former smoker	66 (23.4)	244 (34.0)			
Occasional smoker	11 (3.9)	23 (3.2)			
Heavy smoker	47 (16.7)	80 (11.1)			
<b>Alcohol intake</b>			17.904 (6)	0.133	.007
Never drink	26 (9.2)	110 (15.3)			
Abstinence in the past year	33 (11.7)	120 (16.7)			
Less than once a month	52 (18.4)	147 (20.5)			
About once a month	38 (13.5)	90 (12.5)			
2-4 times a month	84 (29.8)	152 (21.2)			
2-3 times a week	36 (12.8)	69 (9.6)			
4 or more times a week	13 (4.6)	30 (4.2)			

## Feature Selection From Performance-Based Digital Health Literacy

Feature selection using Bayesian linear regression identified 3 features that were statistically associated with digital information comprehension. The number of electronic devices used showed a strong positive association with comprehension scores ( $\beta=.59$ , 95% HDI 0.21-0.93). By contrast, 2 variables related to alcohol intake were negatively associated with comprehension. Specifically, alcohol intake once a month ( $\beta=-1.48$ , 95% HDI  $-2.43$  to 0.49) and alcohol intake 2-4 times a month ( $\beta=-1.05$ , 95% HDI  $-1.76$  to 0.35) were associated with lower performance in digital literacy tasks.

For health care information comprehension, 2 features were statistically significant. Years of education showed a positive association with comprehension performance ( $\beta=.48$ , 95% HDI 0.10-0.88), indicating that higher educational attainment contributes to better understanding of health-related information. Conversely, alcohol intake 2-4 times a month was again associated with reduced performance ( $\beta=-1.03$ , 95% HDI  $-1.96$  to 0.16). Finally, the features selected from the performance-based DHL were the number of electronic devices used, alcohol intake, and years of education. These selected variables were subsequently used in the second-stage predictive modeling of DHL levels.

## Feature Selection From Self-Assessed Digital Health Literacy

The significant predictors associated with 2 DHLS subscales, namely, use of digital devices and health information understanding and utilization, were identified using Bayesian linear regression.

For the digital device utilization subscale, 12 features were statistically significant. Interest in health-related apps was the strongest positive predictor ( $\beta=.44$ , 95% HDI 0.39-0.48). Moreover, several health behavior variables showed positive associations, including abstinence from alcohol intake in the past year ( $\beta=.31$ , 95% HDI 0.09-0.51), alcohol intake about once a month ( $\beta=.37$ , 95% HDI 0.15-0.57), alcohol intake 2-4 times a month ( $\beta=.31$ , 95% HDI 0.11-0.49), alcohol intake 2-3 times a week ( $\beta=.26$ , 95% HDI 0.04-0.52), and being a nonsmoker ( $\beta=.19$ , 95% HDI 0.001-0.37). In addition, self-care confidence ( $\beta=.13$ , 95% HDI 0.03-0.26) was positively associated with digital device use. By contrast, several variables demonstrated significant negative associations with digital device use, including age 70-74 years ( $\beta=-.31$ , 95% HDI  $-0.50$  to 0.11), age 65-69 years ( $\beta=-.17$ , 95% HDI  $-0.31$  to 0.05), number of electronic devices used ( $\beta=-.19$ , 95% HDI  $-0.26$  to 0.13), exercise frequency ( $\beta=-.11$ , 95% HDI  $-0.17$  to 0.05), and years of education ( $\beta=-.07$ , 95% HDI  $-0.14$  to 0.01).

Distributions of digital literacy scores by feature values are presented in [Multimedia Appendix 1](#).

For the subscale on health information understanding and utilization, 7 predictors showed statistically significant associations. Interest in health-related apps emerged as the strongest positive predictor ( $\beta=.34$ , 95% HDI 0.27-0.38). Lack of health information search ( $\beta=.17$ , 95% HDI 0.02-0.32), abstinence from alcohol intake in the past year ( $\beta=.28$ , 95% HDI 0.07-0.55), and former smoking status ( $\beta=.34$ , 95% HDI 0.03-0.69) also showed positive associations with health information understanding and utilization. By contrast, negative associations were observed for the number of electronic devices used ( $\beta=-.14$ , 95% HDI  $-0.21$  to 0.07), age 70-74 years ( $\beta=-.29$ , 95% HDI  $-0.54$  to 0.09), and age 65-69 years ( $\beta=-.25$ , 95% HDI  $-0.41$  to 0.11). These variables, including digital information interest and health behavior factors selected from the self-assessed DHL, were subsequently used as input features in the modeling phase.

## Predictive Modeling of Digital Health Literacy

The performance of the 5 machine learning models was evaluated and compared using multiple classification metrics, including accuracy, precision, and  $F_1$ -score. The detailed results are summarized in [Table 4](#), and the corresponding ROC curves are shown in [Figure 3](#).

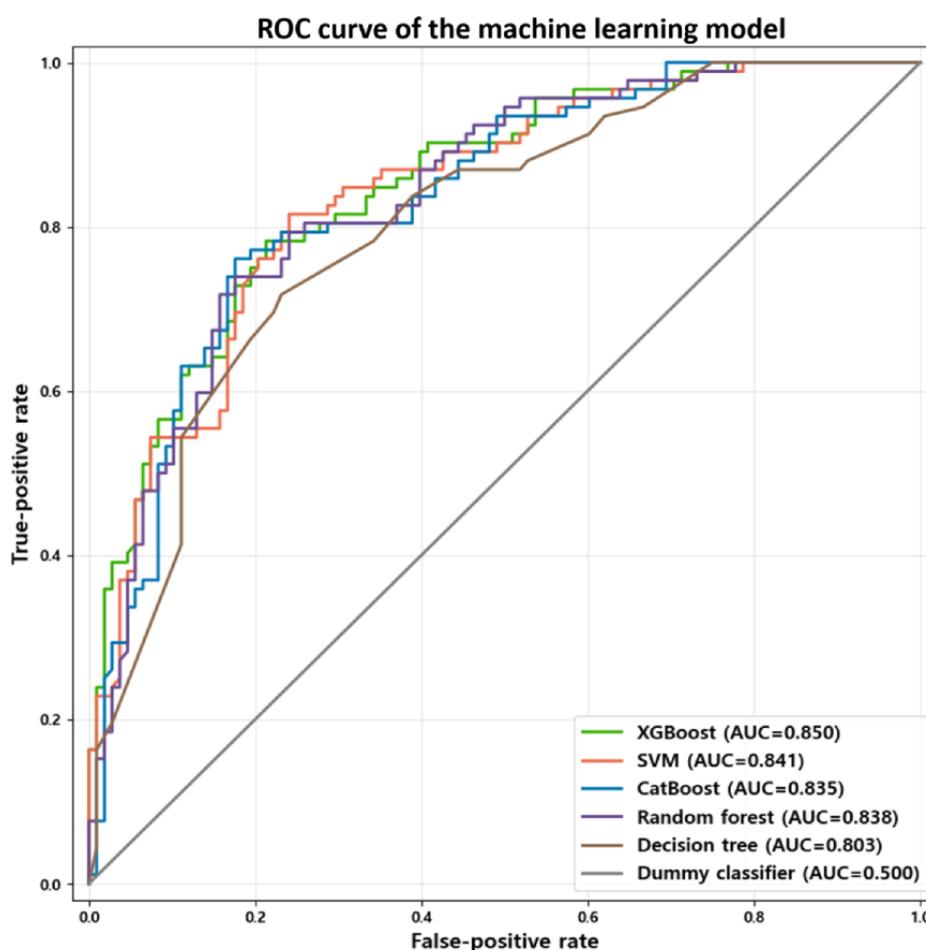
To establish a performance baseline given the class distribution, a dummy classifier was evaluated, yielding an accuracy of 0.540. Relative to this baseline, all machine learning models demonstrated improved predictive performance. Among the 5 models, SVM showed strong results, achieving an accuracy of 0.775, a precision of 0.758, and an  $F_1$ -score of 0.754. By contrast, the decision tree model exhibited the lowest performance across all metrics, with an accuracy of 0.715, a precision of 0.661, and an  $F_1$ -score of 0.716. In addition to standard metrics, ROC curve analysis was conducted to assess the discriminative capability of each model. XGBoost and SVM achieved the highest AUC values (0.850 and 0.841, respectively). Despite these results, CatBoost demonstrated the most balanced and superior overall performance, with an accuracy of 0.785, a precision of 0.769, and an  $F_1$ -score of 0.765. Compared with the dummy classifier, CatBoost improved accuracy by 0.245, which was statistically significant ( $P<.001$ ). This consistently strong performance across primary classification metrics constituted the principal rationale for selecting CatBoost for subsequent SHAP-based post hoc interpretability analysis. These findings indicate that, for DHL prediction, the CatBoost model offers robust and reliable classification performance among the algorithms tested.

**Table 4.** Performance metrics of machine learning models.<sup>a</sup>

Model	Accuracy	Precision	Specificity	Sensitivity	$F_1$ -score	Area under the receiver operating characteristic
Dummy classifier	0.540	0.000	1.000	0.000	0.000	0.500
Decision tree	0.715	0.661	0.657	<i>0.783</i>	0.716	0.803
Support vector machine	0.775	0.758	0.796	0.750	0.754	0.841
Random forest	0.780	<i>0.793</i>	0.843	0.707	0.747	0.838
Extreme gradient boosting	0.780	0.767	0.806	0.750	0.758	<i>0.850</i>
Categorical boosting	<i>0.785</i>	0.769	<i>0.806</i>	0.761	<i>0.765</i>	0.835

<sup>a</sup>Values displayed in italics represent the highest performance achieved by each metric when comparing the performance of the 5 machine learning models.

**Figure 3.** Accuracy and receiver operating characteristic (ROC) curve. AUC: area under the receiver operating characteristic; CatBoost: categorical boosting; SVM: support vector machine; XGBoost: Extreme Gradient Boosting.



### Model Interpretation Using SHAP Analysis

The SHAP analysis was conducted to identify the features that contributed most substantially to the model’s predictions of DHL levels. Figure 4A presents the mean absolute SHAP values for each feature, indicating their relative importance within the model. The 5 most influential features were self-care confidence, lack of health information search, interest in health-related apps, number of electronic devices used, and exercise frequency. These variables played a central role in distinguishing between high- and low-DHL classifications.

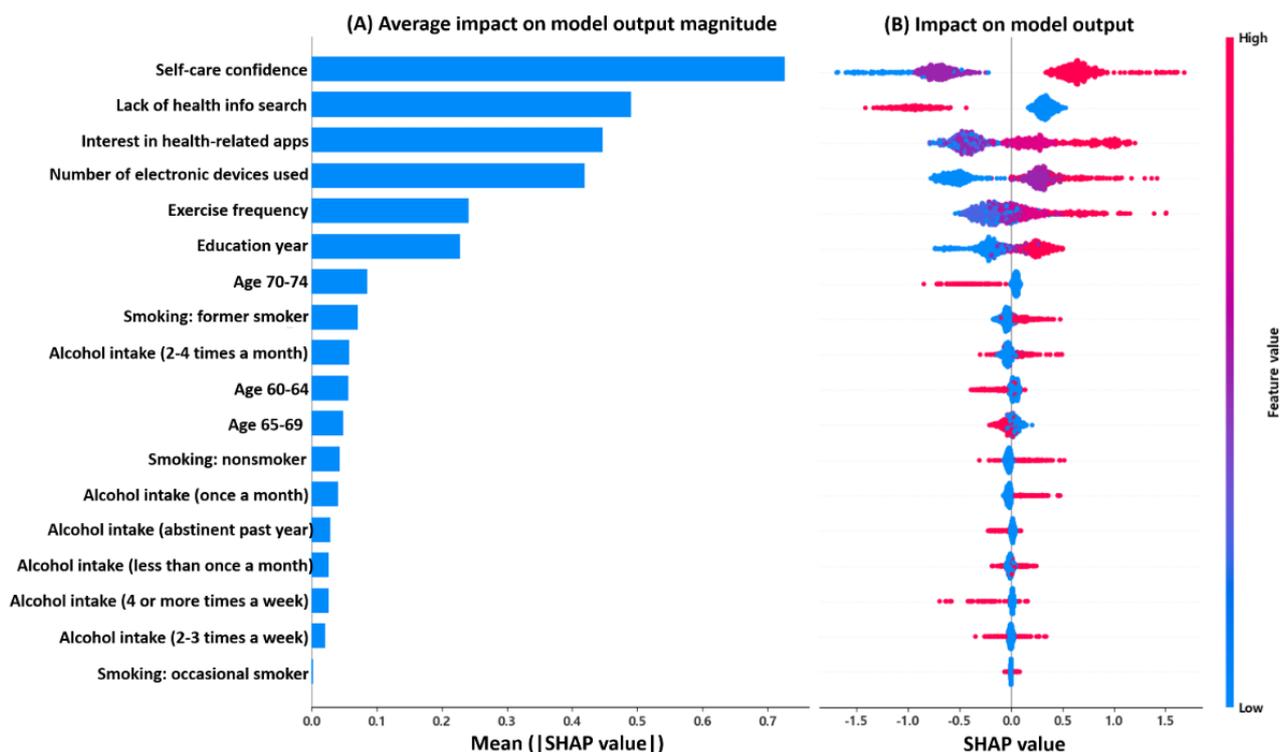
Figure 4B presents a summary plot illustrating the distribution of SHAP values across different feature values. Self-care confidence emerged as the most influential variable, showing the greatest separation between high- and low-DHL predictions. Higher levels of self-care confidence were associated with positive SHAP values, increasing the likelihood of being classified as high DHL. By contrast, lower self-care confidence levels (represented by blue markers) corresponded to negative SHAP values, contributing to low-DHL predictions.

The binary feature representing lack of health information search, derived from one-hot encoding, also demonstrated a

strong effect. Participants who had searched for health information (feature value=0, shown in blue) tended to have positive SHAP values, whereas those who had not (feature value=1, shown in red) exhibited negative SHAP values, indicating a higher likelihood of being classified in the low-DHL group. Similarly, interest in health-related apps and the number of electronic devices used were positively associated with high-DHL predictions. Participants with a greater interest in health-related apps and more frequent use of digital devices showed consistently positive SHAP values, reinforcing their roles as indicators of digital engagement. SHAP values for other

contributing factors, including exercise frequency and education level, were relatively small. Age-related features also had comparatively low influence on model predictions; however, the 70-74 age group consistently appeared in association with lower DHL levels, contributing negatively to the prediction outcomes. Overall, the SHAP analysis provided interpretable evidence of how specific behavioral, psychosocial, and demographic features influenced DHL classification, reinforcing both the explainability and practical utility of the CatBoost model in real-world digital health contexts.

**Figure 4.** Feature importance based on mean Shapley Additive Explanations (SHAP) values according to the contribution of the categorical boosting model to its estimation. (A) Average impact on model output magnitude; (B) Impact on model output.



## Discussion

### Principal Findings

Digital health literacy, defined as the ability to search for, understand, and apply health-related information in digital environments [4], has emerged as a core competency for effectively navigating modern medical and health care services. As health care delivery models evolve from static information provision to dynamic, user-centered platforms, precise assessment of DHL—particularly among older adults—has become increasingly critical. Accurate measurement of DHL is essential for designing tailored interventions that promote equitable access to health care resources and reduce health disparities.

In this study, we developed an explainable machine learning–based predictive framework to evaluate DHL in a large sample of older adults (n=1000). The framework integrates both performance-based and self-assessed DHL measures and incorporates a range of psychosocial and behavioral variables

to identify the most influential predictors contributing to individual differences in digital health literacy.

### Selected Features From Performance-Based Digital Health Literacy

To classify participants into 2 DHL-level groups, explainable features were selected based on the intrinsic evaluation criteria of the Bayesian linear regression model. The analysis revealed that, within the performance-based assessment, key predictors included alcohol intake, number of electronic devices used, and years of education. Among these, the number of electronic devices used emerged as the primary positive predictor of digital information comprehension. This finding is consistent with previous studies showing that older adults who use a wider variety of digital devices are more familiar with online health information and demonstrate stronger skills in information search and utilization [57].

Similarly, years of education positively influenced the comprehension of health information. This finding aligns with previous research identifying educational attainment as a key

social determinant of health literacy. Higher educational attainment is generally associated with a greater capacity to comprehend, critically evaluate, and utilize online health information [16,58]. Individuals with higher levels of education are relatively more advantaged in discerning the credibility of health information and translating it into effective communication with health care providers or self-management behaviors, which in turn facilitates greater acceptance and utilization of digital health services [59,60]. A comparative study involving older adults in Korea and the United States likewise identified educational attainment as a consistent predictor of eHealth literacy, while noting that its impact may be moderated by structural and cultural factors within the digital environment [15]. Collectively, these findings underscore the importance of designing digital health interventions that account for educational disparities and incorporate tailored capacity-building strategies for individuals with lower levels of formal education.

By contrast, participants with higher levels of alcohol consumption exhibited significantly lower performance in both digital and health information comprehension, suggesting that lifestyle factors such as alcohol intake may impair information-processing abilities and negatively affect DHL. Moreover, a previous longitudinal study reported that sustained alcohol consumption is associated with cognitive decline, particularly among older women [61]. Taken together, these results indicate that diversity in digital device usage and higher educational attainment contribute meaningfully to actual information comprehension, and that lifestyle behavior management should be considered a critical component of DHL intervention programs.

### **Selected Features From Self-Assessed Digital Health Literacy**

Self-assessed digital health literacy was primarily influenced by interest in health-related apps, age, alcohol intake, number of electronic devices used, and years of education. Additional contributing factors included self-care confidence, smoking status, and exercise frequency. Specifically, perceived competence in using digital devices was associated with greater confidence in health care decision-making, whereas the ability to understand and apply health information was linked to participants' prior health information-seeking behavior.

Among these variables, interest in health-related apps emerged as the most prominent positive predictor of both DHL subscales. Furthermore, a previous study reported that individuals who frequently engage with health care apps and perceive them as useful or reliable tend to demonstrate stronger competencies in locating, evaluating, and applying health information [57]. Interest in health-related apps could reflect more than mere preference, as it represents a motivational factor that enhances digital engagement. Voluntary use of such apps may foster critical components of DHL, including information searching, selection, appraisal, and application. Therefore, interventions designed to improve DHL should extend beyond digital tools and unidirectional health messaging.

In this study, we found that the older the age group, the lower their tendency to evaluate their understanding of and ability to

use health information. It has been reported that older adults display a reduced ability to adopt digital technology because of cognitive decline [62]. Moreover, previous studies have found that older adults living alone face both cognitive and emotional barriers when adapting to new technologies [63-67]. Addressing these barriers requires tailored educational strategies that combine practical digital training with emotional support.

The number of electronic devices used and years of education proved to be negatively associated with self-assessed DHL. This finding suggests that greater exposure to digital devices or higher educational attainment does not necessarily translate into higher self-assessed DHL. The strength and nature of the relationship between education and digital health literacy may vary depending on the sociocultural context and the health care and digital infrastructure [68]. In addition, individuals with greater actual abilities may apply stricter self-assessment criteria, reflecting more conservative or critical self-assessment tendencies. A prior finding indicated that older adults may experience increased cognitive load when processing complex health-related information, leading to reduced confidence and lower self-ratings [69].

A similar pattern was observed for health information search. In the feature selection stage, lack of health information search was positively associated with the health information understanding and utilization subscale in the self-assessed survey. By contrast, in the model interpretation stage, the same indicator showed negative SHAP values, thereby decreasing the model's predicted likelihood of high DHL. Individuals who do not search for health information may positively assess their understanding of health information even though they lack it, whereas health information seekers might provide cautious self-assessments despite possessing substantial digital health literacy skills. It is known that self-reported responses do not represent a direct readout of an individual's internal state but are rather the product of complex cognitive processes such as comprehension, retrieval, judgment, and response selection, potentially leading to biased or distorted outcomes [70]. These findings emphasize the constraints of self-reported measures and the imperative to validate subjective data with objective performance metrics [71].

In addition, alcohol intake showed a clear divergence by measurement modality. In the self-assessed digital device utilization subscale, alcohol-related categories were positively associated with utilization scores, including abstinence in the past year, alcohol intake about once a month, and alcohol intake 2-4 times a month. By contrast, in the performance-based DHL tasks, alcohol intake once a month and alcohol intake 2-4 times a month were associated with lower task performance. Prior research in older adults indicates that internet use does not consistently correspond to healthier behaviors and that associations may vary by socioeconomic and lifestyle profiles [72]. A population-based study using CHARLS (The China Health and Retirement Longitudinal Study) 2020 data reported that internet use was associated with a lower likelihood of nondrinking, and this association was partially mediated by social participation [73]. This suggests that digital access may increase opportunities for social activities in which alcohol intake is culturally embedded. Taken together,

performance-based DHL tasks are more closely tied to the cognitive and informational demands of digital health activities and may be more sensitive to alcohol-related decrements. By contrast, self-assessed digital device utilization may reflect a broader lifestyle context rather than objective capability alone.

### Development of the Prediction Models

The first stage of the DHL framework identified variables with high explanatory power for both performance-based and self-assessed measures. In the second stage, these selected variables were used to train models predicting DHL levels based on KeHEALS. The model performance evaluation demonstrated that the CatBoost algorithm consistently outperformed the alternative approaches, achieving comparatively high accuracy, precision, and  $F_1$ -scores in predicting DHL among older adults (Table 3). Compared with a conventional single decision tree, CatBoost, which effectively handles categorical variables, proved particularly effective for DHL prediction.

To enhance interpretability, a SHAP analysis was conducted on the CatBoost model, providing empirical evidence of the specific factors influencing DHL (Figure 4). The most influential predictors included self-care confidence, lack of health information search, interest in health-related apps, number of electronic devices used, exercise frequency, and years of education. This confirms that the framework successfully integrates both digital and health literacy-related features into a unified predictive model. Among the key predictors, self-care confidence emerged as the single most important predictor, highlighting the strong association between an individual's confidence in making health-related decisions and their DHL level. This finding aligns with prior research reporting that higher eHealth literacy is significantly associated with greater self-efficacy in older adults [74]. This evidence underscores the potential of DHL-focused interventions to enhance broader health management competencies by strengthening decision-making confidence. Furthermore, consistent with the results of the self-assessed DHL, interest in health-related apps reflected an individual's proactive attitude toward information seeking and app utilization, reinforcing the importance of psychosocial factors alongside technical competencies. Accordingly, interventions aimed at improving DHL should combine technical skill development with psychosocial support to enhance motivation for and engagement with digital health tools [75].

In contrast to the preceding factors, a greater number of electronic devices used, higher exercise frequency, and higher educational attainment were paradoxically associated with lower self-assessed scores for use of digital devices and health information understanding and utilization. These results may reflect the subjective biases inherent in self-reports. Self-report measures are constrained by their susceptibility to respondents' memory and cognitive processes, social desirability and consistency motives, and instrument design features such as item wording, response scales, and question order [76-78]. Alcohol intake showed a similar discrepancy. In the self-assessed digital device utilization subscale, abstinence in the past year and infrequent drinking were positively associated with utilization scores. By contrast, similar frequencies of

alcohol intake were associated with lower performance in the task-based DHL assessment. This pattern suggests that self-assessed utilization may reflect a broader socioeconomic and lifestyle context, whereas performance-based tasks reflect the cognitive and informational demands required for digital health activities.

Therefore, a multimethod approach that integrates both self-assessments and performance-based evaluations is essential for a comprehensive understanding of DHL, rather than relying solely on self-report data. When variables derived from both measures were incorporated into the machine learning model, SHAP analysis indicated that these factors positively contributed to the classification of the high-DHL group, corroborating the results of the performance-based assessment. This demonstrates that reliance on self-reported data alone may lead to biased inferences or the overlooking of important patterns. By combining multifaceted indicators with explainable machine learning techniques, our study's approach enables more precise identification of the determinants of DHL in older adults and provides practical evidence to inform the design of individualized digital health interventions.

KeHEALS is a broadly used and validated assessment tool, although it yields only a single metric for digital literacy. Therefore, it remains uncertain whether individual vulnerabilities stem from fundamental digital device use, comprehension and appraisal of health information, or demographic and psychosocial factors. By using the detailed subscales evaluated within the DHLs, we can elucidate why a particular individual falls into the low-DHL category according to multidimensional criteria. When these domain-informed features are used to predict KeHEALS levels and the resulting predictions are subsequently examined using SHAP, KeHEALS scores can be analytically decomposed into discrete contributions originating from specific digital and health domains, along with additional contextual factors. This study highlights the limitations of single-score screening instruments and advances personalized interventions by offering explanations of feature importance through SHAP analysis.

This study leverages both performance- and survey-based datasets to bridge the gap between self-reported and actual abilities. Using Bayesian regression, the variables explaining each dimension were identified and integrated into the final DHL prediction model. SHAP analysis further validated the consistency between variable importance in the feature-selection stage and the variables' impact on model learning. This dual-phase approach enabled both ante hoc interpretability—by incorporating transparency into the model design from the outset—and post hoc interpretability, addressing the “black box” issue while retaining explanatory power. Finally, the DHL framework offers practical guidance for assessing users' DHL before implementing digital health interventions. This can be particularly valuable in clinical contexts where evaluation and interpretability are critical, thereby supporting the design of tailored and effective health care strategies for older adults.

### Limitations

This study has several limitations that should be noted. First, the performance-based features were identified from a pilot

cohort aged 60-74, whereas the predictive model was trained on a wider age range (55-74 years). Our post hoc comparison revealed that several selected variables differed significantly between the 55-59 and 60-74 age groups, suggesting meaningful demographic and behavioral heterogeneity across age groups. Therefore, the generalizability of the feature-selection process from the small pilot cohort to the younger subgroup may be limited. Future research should employ larger and more diverse samples to enhance the robustness and external validity of the performance-based assessments. Second, because all data collected focused exclusively on older adults in Korea, caution should be exercised when extrapolating the results to populations from different cultural and demographic contexts. Third, the behavioral task was conducted using a single Android smartphone, and all participants were Android users. Thus, familiarity with the Android interface may have been reflected in the digital information comprehension scores and may not fully capture device proficiency among iOS users or across heterogeneous device environments. Fourth, there are inherent limitations related to the measurement scales and variable-processing strategies employed in this study. While offering interpretive simplicity, the binary classification of DHL levels may lead to different outcomes if alternative KeHEALS score thresholds are applied. Furthermore, self-care confidence, which emerged as the primary driver in our predictive model, was assessed using a single-item measure derived from the Korean Health Empowerment Scale. As single-item measures typically exhibit lower reliability and less comprehensive construct coverage than multiitem scales, the high predictive importance of this variable should be interpreted with caution. Additionally, our feature-importance rankings may have been influenced by the mixed encoding strategies used during data preprocessing. Specifically, the use of one-hot encoding for certain lifestyle factors, such as alcohol intake, partitioned these

variables into multiple binary features. This approach may have diluted their aggregate predictive power in the SHAP analysis, potentially leading to an artificial deflation of their importance compared with variables maintained as single ordinal features. Finally, actual DHL is likely to exist along a continuous and multifaceted spectrum. Features that demonstrated unexpected associations in the self-assessed DHL analysis, such as the number of electronic devices used, exercise frequency, and years of education, may reflect the inherent limitations of self-report measures or complex interaction effects with other variables. These observations underscore the need for further in-depth analyses to disentangle these relationships.

## Conclusions

DHL is a key competency for older adults to effectively navigate modern health care. This study established an explainable machine learning framework that integrates performance-based and self-assessed measures to evaluate DHL. Our findings indicate that using a wider variety of electronic devices and having higher educational attainment were positively associated with performance-based DHL, whereas higher alcohol intake was linked to poorer comprehension. Given the limited sample size of the performance-based pilot, these associations warrant caution and require validation in larger cohorts. In the self-assessed domain, DHL levels were primarily related to interest in health-related apps, age, and self-care confidence. In the predictive modeling stage, the CatBoost model demonstrated superior performance, and SHAP analysis identified self-care confidence and interest in health-related apps as key contributors to high DHL. Overall, by combining performance-based and self-assessed measures, this framework offers a structured and interpretable approach to assessing DHL in older adults and can inform the design of tailored digital health interventions in clinical and community settings.

---

## Acknowledgments

This research was supported by the Basic Medical Science Facilitation Program through the Catholic Medical Center of the Catholic University of Korea, funded by the Catholic Education Foundation.

---

## Funding

This research was supported by the Ministry of Science and ICT of Korea and the Korea Medical Device Development Fund (grant RS-2023-00255145).

---

## Data Availability

The data and code underlying this article will be shared on reasonable request to the corresponding author (JWC).

---

## Authors' Contributions

Conceptualization: JWC (lead), CP (equal)  
Data curation: CP  
Formal analysis: CP (lead), SK (supporting)  
Funding acquisition: JWC  
Investigation: YSB, JHK  
Methodology: CP  
Project administration: JWC, JHK  
Validation: JWC, CP  
Visualization: CP (lead), JWC (supporting)

Writing – original draft: CP (lead), JP (supporting)

Writing – review & editing: JWC (lead), CP (supporting), TMK (supporting)

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Distribution of digital literacy and health literacy scores by values of features.

[\[DOCX File , 1139 KB-Multimedia Appendix 1\]](#)

## References

1. Alawiye TR. The impact of digital technology on healthcare delivery and patient outcomes. *ETSN*. 2024;13(02):13-22. [doi: [10.4236/etsn.2024.132002](https://doi.org/10.4236/etsn.2024.132002)]
2. Lim DSC, Kwok B, Williams P, Koczwara B. The impact of digital technology on self-management in cancer: systematic review. *JMIR Cancer*. Nov 22, 2023;9:e45145. [FREE Full text] [doi: [10.2196/45145](https://doi.org/10.2196/45145)] [Medline: [37991831](https://pubmed.ncbi.nlm.nih.gov/37991831/)]
3. Botti A, Monda A. Sustainable value co-creation and digital health: the case of Trentino eHealth Ecosystem. *Sustainability*. Jun 29, 2020;12(13):5263. [doi: [10.3390/su12135263](https://doi.org/10.3390/su12135263)]
4. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res*. Jun 16, 2006;8(2):e9. [FREE Full text] [doi: [10.2196/jmir.8.2.e9](https://doi.org/10.2196/jmir.8.2.e9)] [Medline: [16867972](https://pubmed.ncbi.nlm.nih.gov/16867972/)]
5. Nutakor JA, Zhou L, Larnyo E, Addai-Dansoh S, Cui Y. Impact of health information seeking behavior and digital health literacy on self-perceived health and depression symptoms among older adults in the United States. *Sci Rep*. Dec 28, 2024;14(1):31080. [FREE Full text] [doi: [10.1038/s41598-024-82187-z](https://doi.org/10.1038/s41598-024-82187-z)] [Medline: [39730731](https://pubmed.ncbi.nlm.nih.gov/39730731/)]
6. Ning L, Zheng Z, Liu M, Gao S, Yang X, Bi J, et al. Profiles of digital health literacy among university students and its association with mental health during the COVID-19 pandemic: a latent profile analysis. *Digit Health*. 2024;10:20552076231224596. [FREE Full text] [doi: [10.1177/20552076231224596](https://doi.org/10.1177/20552076231224596)] [Medline: [38205038](https://pubmed.ncbi.nlm.nih.gov/38205038/)]
7. Liu S, Lu Y, Wang D, He X, Ren W, Kong D, et al. Impact of digital health literacy on health-related quality of life in Chinese community-dwelling older adults: the mediating effect of health-promoting lifestyle. *Front Public Health*. 2023;11:1200722. [FREE Full text] [doi: [10.3389/fpubh.2023.1200722](https://doi.org/10.3389/fpubh.2023.1200722)] [Medline: [37415711](https://pubmed.ncbi.nlm.nih.gov/37415711/)]
8. Bíró É, Vincze F, Nagy-Pénczes G, Ádány R. Investigation of the relationship of general and digital health literacy with various health-related outcomes. *Front Public Health*. 2023;11:1229734. [FREE Full text] [doi: [10.3389/fpubh.2023.1229734](https://doi.org/10.3389/fpubh.2023.1229734)] [Medline: [37588120](https://pubmed.ncbi.nlm.nih.gov/37588120/)]
9. The report on the digital divide. National Information Society Agency. 2024. URL: [https://www.nia.or.kr/site/nia\\_kor/ex/bbs/View.do?cbIdx=81623&bcIdx=27832&parentSeq=27832](https://www.nia.or.kr/site/nia_kor/ex/bbs/View.do?cbIdx=81623&bcIdx=27832&parentSeq=27832) [accessed 2025-12-04]
10. Hwang M, Kim G, Lee S, Park Y. Digital health literacy and associated factors among older adults living alone in South Korea: a cross-sectional study. *Res Community Public Health Nurs*. Dec 30, 2024;35:389-400. [doi: [10.12799/rcphn.2024.00766](https://doi.org/10.12799/rcphn.2024.00766)]
11. Shi Z, Du X, Li J, Hou R, Sun J, Marohabutr T. Factors influencing digital health literacy among older adults: a scoping review. *Front Public Health*. 2024;12:1447747. [FREE Full text] [doi: [10.3389/fpubh.2024.1447747](https://doi.org/10.3389/fpubh.2024.1447747)] [Medline: [39555039](https://pubmed.ncbi.nlm.nih.gov/39555039/)]
12. Yang E, Chang SJ, Ryu H, Kim HJ, Jang SJ. Comparing factors associated with eHealth literacy between young and older adults. *J Gerontol Nurs*. Aug 01, 2020;46(8):46-56. [doi: [10.3928/00989134-20200707-02](https://doi.org/10.3928/00989134-20200707-02)] [Medline: [32936926](https://pubmed.ncbi.nlm.nih.gov/32936926/)]
13. Papp-Zipernovszky O, Horváth MD, Schulz PJ, Csabai M. Generation gaps in digital health literacy and their impact on health information seeking behavior and health empowerment in Hungary. *Front Public Health*. 2021;9:635943. [FREE Full text] [doi: [10.3389/fpubh.2021.635943](https://doi.org/10.3389/fpubh.2021.635943)] [Medline: [34055714](https://pubmed.ncbi.nlm.nih.gov/34055714/)]
14. Li S, Cui G, Yin Y, Wang S, Liu X, Chen L. Health-promoting behaviors mediate the relationship between eHealth literacy and health-related quality of life among Chinese older adults: a cross-sectional study. *Qual Life Res*. Aug 2021;30(8):2235-2243. [FREE Full text] [doi: [10.1007/s11136-021-02797-2](https://doi.org/10.1007/s11136-021-02797-2)] [Medline: [33661455](https://pubmed.ncbi.nlm.nih.gov/33661455/)]
15. Lee OE, Kim D, Beum KA. Factors affecting information and communication technology use and eHealth literacy among older adults in the US and South Korea. *Educational Gerontology*. Jul 09, 2020;46(9):575-586. [doi: [10.1080/03601277.2020.1790162](https://doi.org/10.1080/03601277.2020.1790162)]
16. Cherid C, Baghdadli A, Wall M, Mayo NE, Berry G, Harvey EJ, et al. Current level of technology use, health and eHealth literacy in older Canadians with a recent fracture-a survey in orthopedic clinics. *Osteoporos Int*. Jul 2020;31(7):1333-1340. [doi: [10.1007/s00198-020-05359-3](https://doi.org/10.1007/s00198-020-05359-3)] [Medline: [32112118](https://pubmed.ncbi.nlm.nih.gov/32112118/)]
17. Berkowsky RW. Exploring predictors of eHealth literacy among older adults: findings from the 2020 CALSPEAKS Survey. *Gerontol Geriatr Med*. 2021;7:23337214211064227. [FREE Full text] [doi: [10.1177/23337214211064227](https://doi.org/10.1177/23337214211064227)] [Medline: [34926723](https://pubmed.ncbi.nlm.nih.gov/34926723/)]
18. Quinn S, Bond R, Nugent C. Quantifying health literacy and eHealth literacy using existing instruments and browser-based software for tracking online health information seeking behavior. *Computers in Human Behavior*. Apr 2017;69:256-267. [FREE Full text] [doi: [10.1016/j.chb.2016.12.032](https://doi.org/10.1016/j.chb.2016.12.032)]

19. Komalavalli D, Lopson S, Karthik M. Integrated prediction of mental health and fitness using machine learning models. New York, NY. IEEE; 2024. Presented at: 2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS); December 17-18, 2024:557-562; Bengaluru, Karnataka, India. [doi: [10.1109/ICICNIS64247.2024.10823203](https://doi.org/10.1109/ICICNIS64247.2024.10823203)]
20. Lundberg S, Lee S. A unified approach to interpreting model predictions. New York, NY. Curran Associates Inc; 2017. Presented at: The 31st International Conference on Neural Information Processing Systems (NIPS'17); December 4-9, 2017:4768-4777; Long Beach, CA. [doi: [10.5555/3295222.3295230](https://doi.org/10.5555/3295222.3295230)]
21. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. *J Med Internet Res*. Nov 14, 2006;8(4):e27. [FREE Full text] [doi: [10.2196/jmir.8.4.e27](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
22. Thorup CB, Uitto M, Butler-Henderson K, Wamala-Andersson S, Hoffrén-Mikkola M, Schack Thoft D, et al. Choosing the best digital health literacy measure for research: mixed methods study. *J Med Internet Res*. Apr 08, 2025;27:e59807. [FREE Full text] [doi: [10.2196/59807](https://doi.org/10.2196/59807)] [Medline: [40198098](https://pubmed.ncbi.nlm.nih.gov/40198098/)]
23. Norman C. eHealth literacy 2.0: problems and opportunities with an evolving concept. *J Med Internet Res*. Dec 23, 2011;13(4):e125. [FREE Full text] [doi: [10.2196/jmir.2035](https://doi.org/10.2196/jmir.2035)] [Medline: [22193243](https://pubmed.ncbi.nlm.nih.gov/22193243/)]
24. Chaniaud N, Sagnier C, Loup-Escande E. Translation and validation study of the French version of the eHealth Literacy Scale: web-based survey on a student population. *JMIR Form Res*. Aug 31, 2022;6(8):e36777. [FREE Full text] [doi: [10.2196/36777](https://doi.org/10.2196/36777)] [Medline: [36044264](https://pubmed.ncbi.nlm.nih.gov/36044264/)]
25. Xu RH, Zhou L, Lu SY, Wong EL, Chang J, Wang D. Psychometric validation and cultural adaptation of the Simplified Chinese eHealth Literacy Scale: cross-sectional study. *J Med Internet Res*. Dec 07, 2020;22(12):e18613. [FREE Full text] [doi: [10.2196/18613](https://doi.org/10.2196/18613)] [Medline: [33284123](https://pubmed.ncbi.nlm.nih.gov/33284123/)]
26. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. *J Med Internet Res*. Nov 09, 2011;13(4):e86. [FREE Full text] [doi: [10.2196/jmir.1840](https://doi.org/10.2196/jmir.1840)] [Medline: [22071338](https://pubmed.ncbi.nlm.nih.gov/22071338/)]
27. Petrič G, Atanasova S. Validation of the extended e-health literacy scale: structural validity, construct validity and measurement invariance. *BMC Public Health*. Jul 25, 2024;24(1):1991. [FREE Full text] [doi: [10.1186/s12889-024-19431-8](https://doi.org/10.1186/s12889-024-19431-8)] [Medline: [39054541](https://pubmed.ncbi.nlm.nih.gov/39054541/)]
28. Marsall M, Engelmann G, Skoda E, Teufel M, Bäuerle A. Measuring electronic health literacy: development, validation, and test of measurement invariance of a revised German Version of the eHealth Literacy Scale. *J Med Internet Res*. Feb 02, 2022;24(2):e28252. [FREE Full text] [doi: [10.2196/28252](https://doi.org/10.2196/28252)] [Medline: [35107437](https://pubmed.ncbi.nlm.nih.gov/35107437/)]
29. Luz S, Nogueira P, Costa A, Henriques A. Psychometric analysis of the eHealth Literacy Scale in Portuguese older adults (eHEALS-PT24): instrument development and validation. *J Med Internet Res*. Feb 26, 2025;27:e57730. [FREE Full text] [doi: [10.2196/57730](https://doi.org/10.2196/57730)] [Medline: [40009846](https://pubmed.ncbi.nlm.nih.gov/40009846/)]
30. Lin C, Broström A, Griffiths MD, Pakpour AH. Psychometric evaluation of the Persian eHealth Literacy Scale (eHEALS) among elder Iranians with heart failure. *Eval Health Prof*. Dec 2020;43(4):222-229. [doi: [10.1177/0163278719827997](https://doi.org/10.1177/0163278719827997)] [Medline: [30744419](https://pubmed.ncbi.nlm.nih.gov/30744419/)]
31. Kim H, Yang E, Ryu H, Kim HJ, Jang SJ, Chang SJ. Psychometric comparisons of measures of eHealth literacy using a sample of Korean older adults. *Int J Older People Nurs*. May 2021;16(3):e12369. [doi: [10.1111/ohn.12369](https://doi.org/10.1111/ohn.12369)] [Medline: [33527701](https://pubmed.ncbi.nlm.nih.gov/33527701/)]
32. Hernández Encuentra E, Robles N, Angulo-Brunet A, Cullen D, Del Arco I. Spanish and Catalan versions of the eHealth Literacy Questionnaire: translation, cross-cultural adaptation, and validation study. *J Med Internet Res*. May 10, 2024;26:e49227. [FREE Full text] [doi: [10.2196/49227](https://doi.org/10.2196/49227)] [Medline: [38728072](https://pubmed.ncbi.nlm.nih.gov/38728072/)]
33. Chung S, Nahm E. Testing reliability and validity of the eHealth Literacy Scale (eHEALS) for older adults recruited online. *Comput Inform Nurs*. Apr 2015;33(4):150-156. [FREE Full text] [doi: [10.1097/CIN.000000000000146](https://doi.org/10.1097/CIN.000000000000146)] [Medline: [25783223](https://pubmed.ncbi.nlm.nih.gov/25783223/)]
34. Kyaw MY, Aung MN, Koyanagi Y, Moolphate S, Aung TNN, Ma HKC, et al. Sociodigital determinants of eHealth literacy and related impact on health outcomes and eHealth use in Korean older adults: community-based cross-sectional survey. *JMIR Aging*. Aug 13, 2024;7:e56061. [FREE Full text] [doi: [10.2196/56061](https://doi.org/10.2196/56061)] [Medline: [39140239](https://pubmed.ncbi.nlm.nih.gov/39140239/)]
35. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: measuring a broad spectrum of Health 1.0 and Health 2.0 skills. *J Med Internet Res*. Jan 24, 2017;19(1):e27. [doi: [10.2196/jmir.6709](https://doi.org/10.2196/jmir.6709)] [Medline: [28119275](https://pubmed.ncbi.nlm.nih.gov/28119275/)]
36. Kim S, Park C, Park S, Kim D, Bae Y, Kang J, et al. Measuring digital health literacy in older adults: development and validation study. *J Med Internet Res*. Feb 05, 2025;27:e65492. [FREE Full text] [doi: [10.2196/65492](https://doi.org/10.2196/65492)] [Medline: [39908081](https://pubmed.ncbi.nlm.nih.gov/39908081/)]
37. Zhao B, Huang L, Cheng X, Chen T, Li S, Wang X, et al. Digital health literacy and associated factors among internet users from China: a cross-sectional study. *BMC Public Health*. Mar 27, 2024;24(1):908. [FREE Full text] [doi: [10.1186/s12889-024-18324-0](https://doi.org/10.1186/s12889-024-18324-0)] [Medline: [38539176](https://pubmed.ncbi.nlm.nih.gov/38539176/)]
38. Liu S, Zhao H, Fu J, Kong D, Zhong Z, Hong Y, et al. Current status and influencing factors of digital health literacy among community-dwelling older adults in Southwest China: a cross-sectional study. *BMC Public Health*. May 17, 2022;22(1):996. [FREE Full text] [doi: [10.1186/s12889-022-13378-4](https://doi.org/10.1186/s12889-022-13378-4)] [Medline: [35581565](https://pubmed.ncbi.nlm.nih.gov/35581565/)]

39. Ezeigwe OJ, Nwosu KOS, Afolayan OK, Ojaruega AA, Echere J, Desai M, et al. Technological-based interventions in cancer and factors associated with the use of mobile digital wellness and health apps among cancer information seekers: cross-sectional study. *J Med Internet Res*. Feb 05, 2025;27:e63403. [doi: [10.2196/63403](https://doi.org/10.2196/63403)] [Medline: [39909418](https://pubmed.ncbi.nlm.nih.gov/39909418/)]
40. Dunning D, Heath C, Suls JM. Flawed self-assessment: implications for health, education, and the workplace. *Psychol Sci Public Interest*. Dec 2004;5(3):69-106. [doi: [10.1111/j.1529-1006.2004.00018.x](https://doi.org/10.1111/j.1529-1006.2004.00018.x)] [Medline: [26158995](https://pubmed.ncbi.nlm.nih.gov/26158995/)]
41. Vowels MJ. Trying to outrun causality with machine learning: limitations of model explainability techniques for exploratory research. *Psychol Methods*. Sep 09, 2024:met0000699. (forthcoming). [doi: [10.1037/met0000699](https://doi.org/10.1037/met0000699)] [Medline: [39250291](https://pubmed.ncbi.nlm.nih.gov/39250291/)]
42. Bonovas S, Piovani D. Simpson's paradox in clinical research: a cautionary tale. *J Clin Med*. Feb 18, 2023;12(4):1633. [[FREE Full text](#)] [doi: [10.3390/jcm12041633](https://doi.org/10.3390/jcm12041633)] [Medline: [36836181](https://pubmed.ncbi.nlm.nih.gov/36836181/)]
43. Dwyer DB, Falkai P, Koutsouleris N. Machine learning approaches for clinical psychology and psychiatry. *Annu Rev Clin Psychol*. May 07, 2018;14:91-118. [doi: [10.1146/annurev-clinpsy-032816-045037](https://doi.org/10.1146/annurev-clinpsy-032816-045037)] [Medline: [29401044](https://pubmed.ncbi.nlm.nih.gov/29401044/)]
44. İnceoğlu F, Deniz S, Yagin FH. Prediction of effective sociodemographic variables in modeling health literacy: a machine learning approach. *Int J Med Inform*. Oct 2023;178:105167. [doi: [10.1016/j.ijmedinf.2023.105167](https://doi.org/10.1016/j.ijmedinf.2023.105167)] [Medline: [37572386](https://pubmed.ncbi.nlm.nih.gov/37572386/)]
45. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. Jan 2020;2(1):56-67. [[FREE Full text](#)] [doi: [10.1038/s42256-019-0138-9](https://doi.org/10.1038/s42256-019-0138-9)] [Medline: [32607472](https://pubmed.ncbi.nlm.nih.gov/32607472/)]
46. Chakraborty C, Bhattacharya M, Pal S, Lee S. From machine learning to deep learning: advances of the recent data-driven paradigm shift in medicine and healthcare. *Current Research in Biotechnology*. 2024;7:100164. [doi: [10.1016/j.crbiot.2023.100164](https://doi.org/10.1016/j.crbiot.2023.100164)]
47. Alhumaidi NH, Dermawan D, Kamaruzaman HF, Alotaqi N. The use of machine learning for analyzing real-world data in disease prediction and management: systematic review. *JMIR Med Inform*. Jun 19, 2025;13:e68898. [[FREE Full text](#)] [doi: [10.2196/68898](https://doi.org/10.2196/68898)] [Medline: [40537090](https://pubmed.ncbi.nlm.nih.gov/40537090/)]
48. Chang SJ, Yang E, Ryu H, Kim HJ, Yoon JY. Cross-cultural adaptation and validation of the eHealth Literacy Scale in Korea. *Korean J Adult Nurs*. 2018;30(5):504-515. [doi: [10.7475/kjan.2018.30.5.504](https://doi.org/10.7475/kjan.2018.30.5.504)]
49. Park C, Park Y. Validity and reliability of Korean version of Health Empowerment Scale (K-HES) for older adults. *Asian Nurs Res (Korean Soc Nurs Sci)*. Sep 2013;7(3):142-148. [[FREE Full text](#)] [doi: [10.1016/j.anr.2013.07.004](https://doi.org/10.1016/j.anr.2013.07.004)] [Medline: [25030252](https://pubmed.ncbi.nlm.nih.gov/25030252/)]
50. Baldwin SA, Larson MJ. An introduction to using Bayesian linear regression with clinical data. *Behav Res Ther*. Nov 2017;98:58-75. [doi: [10.1016/j.brat.2016.12.016](https://doi.org/10.1016/j.brat.2016.12.016)] [Medline: [28081861](https://pubmed.ncbi.nlm.nih.gov/28081861/)]
51. Quinlan J. Induction of decision trees. *Mach Learn*. Mar 1986;1(1):81-106. [doi: [10.1023/A:1022643204877](https://doi.org/10.1023/A:1022643204877)]
52. Prokhorenkova L, Gusev G, Vorobev A, Dorogush A, Gulin A. CatBoost: unbiased boosting with categorical features. Red Hook, NY: Curran Associates Inc; 2018. Presented at: The 32nd International Conference on Neural Information Processing Systems (NIPS'18); December 3-8, 2018:6639-6649; Montréal, QC, Canada. [doi: [10.5555/3327757.3327770](https://doi.org/10.5555/3327757.3327770)]
53. Cortes C, Vapnik V. Support-vector networks. *Machine Learning*. Sep 21, 1995;20(3):273-297. [doi: [10.1023/A:1022627411411](https://doi.org/10.1023/A:1022627411411)]
54. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. New York, NY: Association for Computing Machinery; 2016. Presented at: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16); August 13-17, 2016:785-794; San Francisco, CA. [doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)]
55. Breiman L. Random forests. *Machine Learning*. Oct 21, 2001;45(1):5-32. [doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]
56. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: a next-generation hyperparameter optimization framework. New York, NY: Association for Computing Machinery; 2019. Presented at: The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD'19); August 4-8, 2019:2623-2631; Anchorage, AK. [doi: [10.1145/3292500.3330701](https://doi.org/10.1145/3292500.3330701)]
57. Tennant B, Stelfox M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. *J Med Internet Res*. Mar 17, 2015;17(3):e70. [[FREE Full text](#)] [doi: [10.2196/jmir.3992](https://doi.org/10.2196/jmir.3992)] [Medline: [25783036](https://pubmed.ncbi.nlm.nih.gov/25783036/)]
58. Green G. Seniors' eHealth literacy, health and education status and personal health knowledge. *Digit Health*. 2022;8:20552076221089803. [[FREE Full text](#)] [doi: [10.1177/20552076221089803](https://doi.org/10.1177/20552076221089803)] [Medline: [35371531](https://pubmed.ncbi.nlm.nih.gov/35371531/)]
59. Seckin G, Hughes S. Patient-reported outcomes in a nationally representative sample of older internet users: cross-sectional survey. *JMIR Aging*. Nov 24, 2021;4(4):e16006. [[FREE Full text](#)] [doi: [10.2196/16006](https://doi.org/10.2196/16006)] [Medline: [34822340](https://pubmed.ncbi.nlm.nih.gov/34822340/)]
60. Price-Haywood EG, Harden-Barrios J, Ulep R, Luo Q. eHealth literacy: patient engagement in identifying strategies to encourage use of patient portals among older adults. *Popul Health Manag*. Dec 2017;20(6):486-494. [doi: [10.1089/pop.2016.0164](https://doi.org/10.1089/pop.2016.0164)] [Medline: [28384076](https://pubmed.ncbi.nlm.nih.gov/28384076/)]
61. Han L, Jia J. Alcohol consumption, poor lifestyle choices, and air pollution worsen cognitive function in seniors: a cohort study in China. *Environ Sci Pollut Res Int*. Apr 2022;29(18):26877-26888. [doi: [10.1007/s11356-021-17891-8](https://doi.org/10.1007/s11356-021-17891-8)] [Medline: [34860343](https://pubmed.ncbi.nlm.nih.gov/34860343/)]
62. Wan X, Lighthall NR, Xie R. Consistent and robust predictors of internet use among older adults over time identified by machine learning. *Computers in Human Behavior*. Dec 2022;137:107413. [doi: [10.1016/j.chb.2022.107413](https://doi.org/10.1016/j.chb.2022.107413)]

63. Charness N, Boot WR. Aging and information technology use. *Curr Dir Psychol Sci.* Oct 01, 2009;18(5):253-258. [doi: [10.1111/j.1467-8721.2009.01647.x](https://doi.org/10.1111/j.1467-8721.2009.01647.x)]
64. Lee C, Coughlin JF. PERSPECTIVE: Older adults' adoption of technology: an integrated approach to identifying determinants and barriers. *J Product Innov Manag.* Jun 03, 2014;32(5):747-759. [doi: [10.1111/jpim.12176](https://doi.org/10.1111/jpim.12176)]
65. Peek STM, Luijkx KG, Rijnaard MD, Nieboer ME, van der Voort CS, Aarts S, et al. Older adults' reasons for using technology while aging in place. *Gerontology.* 2016;62(2):226-237. [FREE Full text] [doi: [10.1159/000430949](https://doi.org/10.1159/000430949)] [Medline: [26044243](https://pubmed.ncbi.nlm.nih.gov/26044243/)]
66. Piper AM, Brewer R, Cornejo R. Technology learning and use among older adults with late-life vision impairments. *Univ Access Inf Soc.* Oct 21, 2016;16(3):699-711. [doi: [10.1007/s10209-016-0500-1](https://doi.org/10.1007/s10209-016-0500-1)]
67. Schroeder T, Dodds L, Georgiou A, Gewalt H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. *JMIR Aging.* May 16, 2023;6:e44564. [FREE Full text] [doi: [10.2196/44564](https://doi.org/10.2196/44564)] [Medline: [37191976](https://pubmed.ncbi.nlm.nih.gov/37191976/)]
68. Yuen E, Winter N, Savira F, Huggins CE, Nguyen L, Cooper P, et al. Digital health literacy and its association with sociodemographic characteristics, health resource use, and health outcomes: rapid review. *Interact J Med Res.* Jul 26, 2024;13:e46888. [FREE Full text] [doi: [10.2196/46888](https://doi.org/10.2196/46888)] [Medline: [39059006](https://pubmed.ncbi.nlm.nih.gov/39059006/)]
69. Vroman KG, Arthanat S, Lysack C. "Who over 65 is online?" Older adults' dispositions toward information communication technology. *Computers in Human Behavior.* Feb 2015;43:156-166. [doi: [10.1016/j.chb.2014.10.018](https://doi.org/10.1016/j.chb.2014.10.018)]
70. Schwarz N. Self-reports: how the questions shape the answers. *American Psychologist.* Feb 1999;54(2):93-105. [doi: [10.1037/0003-066x.54.2.93](https://doi.org/10.1037/0003-066x.54.2.93)]
71. Kaye S, Lewis I, Freeman J. Comparison of self-report and objective measures of driving behavior and road safety: a systematic review. *J Safety Res.* Jun 2018;65:141-151. [doi: [10.1016/j.jsr.2018.02.012](https://doi.org/10.1016/j.jsr.2018.02.012)] [Medline: [29776523](https://pubmed.ncbi.nlm.nih.gov/29776523/)]
72. Duplaga M. The association between internet use and health-related outcomes in older adults and the elderly: a cross-sectional study. *BMC Med Inform Decis Mak.* May 06, 2021;21(1):150. [FREE Full text] [doi: [10.1186/s12911-021-01500-2](https://doi.org/10.1186/s12911-021-01500-2)] [Medline: [33957909](https://pubmed.ncbi.nlm.nih.gov/33957909/)]
73. Yang X, Chen Y, Qiu T, Zhu X. The mediation effect of social participation in the relationship between internet use and health behaviors among middle-aged and older individuals. *Digit Health.* 2025;11:20552076251346239. [FREE Full text] [doi: [10.1177/20552076251346239](https://doi.org/10.1177/20552076251346239)] [Medline: [40438191](https://pubmed.ncbi.nlm.nih.gov/40438191/)]
74. Park C. Electronic health literacy as a source of self-efficacy among community-dwelling older adults. *Clin Gerontol.* Jul 01, 2024:1-10. [FREE Full text] [doi: [10.1080/07317115.2024.2373894](https://doi.org/10.1080/07317115.2024.2373894)] [Medline: [38949203](https://pubmed.ncbi.nlm.nih.gov/38949203/)]
75. Fan S, Jain RC, Kankanhalli MS. A comprehensive picture of factors affecting user willingness to use mobile health applications. *ACM Trans Comput Healthcare.* Jan 13, 2024;5(1):1-31. [doi: [10.1145/3626962](https://doi.org/10.1145/3626962)]
76. Podsakoff PM, MacKenzie SB, Lee J, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol.* Oct 2003;88(5):879-903. [doi: [10.1037/0021-9010.88.5.879](https://doi.org/10.1037/0021-9010.88.5.879)] [Medline: [14516251](https://pubmed.ncbi.nlm.nih.gov/14516251/)]
77. Huprich SK, Bornstein RF, Schmitt TA. Self-report methodology is insufficient for improving the assessment and classification of axis II personality disorders. *J Pers Disord.* Oct 2011;25(5):557-570. [doi: [10.1521/pedi.2011.25.5.557](https://doi.org/10.1521/pedi.2011.25.5.557)] [Medline: [22023295](https://pubmed.ncbi.nlm.nih.gov/22023295/)]
78. Haefel GJ, Howard GS. Self-report: psychology's four-letter word. *Am J Psychol.* 2010;123(2):181-188. [doi: [10.5406/amerjpsyc.123.2.0181](https://doi.org/10.5406/amerjpsyc.123.2.0181)] [Medline: [20518434](https://pubmed.ncbi.nlm.nih.gov/20518434/)]

## Abbreviations

- AUC:** area under the receiver operating characteristic
- CatBoost:** categorical boosting
- CHARLS:** The China Health and Retirement Longitudinal Study
- DHL:** digital health literacy
- DHLS:** Digital Health Literacy Scale
- eHEALS:** eHealth Literacy Scale
- HDI:** highest density interval
- KeHEALS:** Korean version of the eHealth Literacy Scale
- mHealth:** mobile health
- ROC:** receiver operating characteristic
- SHAP:** Shapley Additive Explanations
- SVM:** support vector machine
- XGBoost:** Extreme Gradient Boosting

*Edited by A Coristine; submitted 20.Oct.2025; peer-reviewed by F Fischer; comments to author 17.Nov.2025; revised version received 25.Jan.2026; accepted 26.Jan.2026; published 23.Mar.2026*

*Please cite as:*

*Park C, Park J, Kim S, Bae YS, Kang J-H, Kim T-M, Chun J-W*

*Explainable Machine Learning for Assessing Digital Health Literacy in Older Adults: Validation and Development of a Two-Stage Model Integrating Performance-Based and Self-Assessed Indicators*

*JMIR Med Inform 2026;14:e86171*

*URL: <https://medinform.jmir.org/2026/1/e86171>*

*doi: [10.2196/86171](https://doi.org/10.2196/86171)*

*PMID:*

©Choonghee Park, Jiyeon Park, Seora Kim, Ye Seul Bae, Jae-Heon Kang, Tae-Min Kim, Ji-Won Chun. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 23.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.