ELSEVIER
Contents lists available at ScienceDirect
Journal of Biomedical Informatics
journal homepage: www.elsevier.com/locate/yjbin

Modelling Assistive Technology Adoption for People with Dementia

Priyanka Chaurasia [a*], Sally I. McClean [b], Chris D. Nugent [a], Ian Cleland [a], Shuai Zhang [a], Mark P. Donnelly [a], Bryan W. Scotney [b], Chelsea Sanders [c], Ken Smith [d], Maria C. Norton [c], JoAnn Tschanz [c]

[a] *School of Computing and Mathematics, University of Ulster, Newtownabbey, Co. Antrim, Northern Ireland, BT37 0QB.*
[b] *School of Computing and Information Engineering, University of Ulster, Coleraine, Northern Ireland, BT52 1SA.*
[c] *Department of Psychology, Utah State University, Logan, UT 84322-4440, USA.*
[d] *Population Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA.*
[e] *Department of Family, Consumer, and Human Development, Utah State University, Logan, UT 84322-4440, USA.*

ABSTRACT

Purpose: Assistive technologies have been identified as a potential solution for the provision of elderly care. Such technologies have in general the capacity to enhance the quality of life and increase the level of independence among their users. Nevertheless, the acceptance of these technologies is crucial to their success. Generally speaking, the elderly are not well-disposed to technologies and have limited experience; these factors contribute towards limiting the widespread acceptance of technology. It is therefore important to evaluate the potential success of technologies prior to their deployment.

Materials and methods: The research described in this paper builds upon our previous work on modeling adoption of assistive technology, in the form of cognitive prosthetics such as reminder apps and aims at identifying a refined sub-set of features which offer improved accuracy in predicting technology adoption. Consequently, in this paper, an adoption model is built using a set of features extracted from a user's background to minimise the likelihood of non-adoption. The work is based on analysis of data from the Cache County Study on Memory and Aging (CCSMA) with 31 features covering a range of age, gender, education and details of health condition. In the process of modelling adoption, feature selection and feature reduction is carried out followed by identifying the best classification models.

Findings: With the reduced set of labelled features the technology adoption model built achieved an average prediction accuracy of 92.48% when tested on 173 participants.

Conclusions: We conclude that modelling user adoption from a range of parameters such as physical, environmental and social perspectives is beneficial in recommending a technology to a particular user based on their profile.

* Corresponding author: p.chaurasia@ulster.ac.uk (Priyanka Chaurasia).

1. Introduction

In order to address the rising demands for care of older patients with early stage memory impairments, a new area of research has emerged that aims to support people through the use of technology-based assistive solutions [1], [2], [3], [4], [5]. With the goal of offering a means of independent living, assistive technologies bring intelligence to the surroundings and proactively support users with daily activities [6]. One specific solution relates to tools which can support the cognitive functioning of the elderly. These tools range from audio systems that, for example, guide users, step-by-step, through the process of washing their hands to monitoring devices that can alert caregivers or healthcare professionals when an individual is at risk of falling [7]. These tools are, however, beneficial only if they are fully adopted by the user. The cognitively impaired population is comprised of semi-autonomous persons, some of which require some form of assistance to assist in independently completing activities with minor levels of supervision [7].

A key requirement of a technology-based assistive solution is to understand the normal or expected lifestyle of each individual and consequently suggest relevant solutions based on their willingness to adopt the technology and their current cognitive abilities [7]. Understanding the factors that affect technology adoption is significant for the long term success of these assistive technologies. Incorporation of such information in learning adoption decisions has proved to be beneficial and is a relatively new area of research [8], [9]. A prior assessment of the likelihood of adoption, through prediction models, before the technology is introduced to people with dementia (PwD) can avoid disappointment and waste of resources. Prior assessment can also save time, make caregiving easier logistically, make the care recipient feel safer, increase feelings of being effective and reduce stress by basing deployment decisions on the prediction of adoption success. Additionally, if the model successfully identifies a PwD's unsuitability to such technologies, the associated cost of not using a solution can be minimized and the resulting negative impact on the quality of life of a PwD can be avoided by providing alternative solutions such as caregiver support or behavioural therapy. Previous efforts have largely been directed towards developing technologies with the notion of technology adoption largely being overlooked. Recent research has evidenced that there are benefits in considering the profile of a potential user when considering introducing a form of assistive technology [8], [9].

The focus of our work involves the analysis of a user's background prior to recommending a solution. Our work is based on analysis of data from the CCSMA along with information relating to an individual's compliance with the usage of a reminding application. We also have an additional medical data from the Utah Population Database (UPDB) but in the current work described in this paper we have only used the self-proclaimed medical data held in CCSMA, which is easier to obtain since it comes directly from the patient. In the future work our aim would be to combine all these datasets. These information would be considered in conjunction with details of cognitive assessments from the CCSMA and medical and genealogical related details from the UPDB and will subsequently forms the basis of the inputs to the development of a new technology adoption model. The vision of developing the model is to build a predictive tool that can be used by healthcare professionals in determining the suitability of a certain technology for a PwD.

In order to address the above problems, we developed technology adoption models based on the patient self-proclaimed information and medical data that are gathered during consultation with the PwDs and their carers, available in the CCSMA and UPDB datasets, respectively. Our previous research in the area of technology adoption models aimed to characterise individuals with dementia and identify features that may be relevant in the determination of the adoption of assistive technology [8], [9]. Features were collected through an iterative design process, involving evaluations with 40 participants with dementia. Features included age, gender, mini mental state exam (MMSE) score, profession, technology, experience and environmental conditions such as access to broadband, mobile reception and living arrangement. Based on these collected features, optimal predictive models were developed. Overall, the model trained using the *k*-nearest neighbour (*k*NN) classification algorithm using the features gender, living arrangement, MMSE, broadband, age, mobile reception, and caregiver gave performance levels in the region of 84% in terms of predicting those who would adopt an assistive technology. Nevertheless, it was noted that the prediction models may have been limited by the size of the training data.

The current work described in this paper builds upon our previous work on modeling adoption and aims at identifying a refined sub-set of features which offer improved accuracy in predicting technology adoption based on the CCSMA dataset. Keeping the above discussion in mind and building upon our previous research, this paper introduces the findings from the Technology Adoption and Usage Tool (TAUT). The TAUT project engaged with PwD associated with the CCSMA. Each participant was enrolled on a 12 month evaluation study of the TAUT reminder application [10]. This system was designed by a multidisciplinary team through an iterative design process and has been previously evaluated on a small scale with a representative cohort [2]. The current version of the solution, described in [10], was developed for the android and iPhone platform and was designed to provide the user with an interface to schedule and acknowledge reminders for a range of activities daily living (ADL) including, medication, meals, appointments and bathing. The reminders can be set by the PwD, or by a caregiver or family member and are delivered at a specified time and are presented as a popup dialog box on the mobile's screen accompanied by a picture indicating the type of ADL, a textual description and a melodic tone. The work presented here describes how the set of relevant features that would affect adoption with the ambition of building highly accepted, well-designed, functional and market-oriented assistive technologies was identified. Our contribution to knowledge being presented in the paper can be considered as the following:

1. identifying features that are directly related to adoption.
2. identifying features that are indirectly related to adoption.
3. successfully modelling adoption using a range of classification algorithms.

The work described in this paper analyses and interprets the data available from the CCSMA dataset [11] for the adoption of the TAUT reminder application. In the process of recruiting participants from the CCSMA for the study, we report features that are more likely to affect adoption. The aim is to develop predictive models with a set of features that can be easily gathered. As a benchmark for comparison, we compare our results with the results obtained from our previous studies. The remainder of the paper is organised as follows: the related literature is discussed in Section 2. Section 3 provides details of the data and previous work carried out. Section 4 details the proposed methodology, approach followed and evaluation carried out. Finally, Section 5 provides the conclusion to the work and describes the problems which will be considered in future work.

2. Related work

The acceptance of technology is a critical issue and efforts have been made in previous research to investigate the factors that could hinder the successful deployment of assistive solutions in elderly care [12]. To predict technology adoption, models such as the technology acceptance model (TAM) [13], [3] have been developed in the past. TAM identifies that the user behaviour is influenced by the perceived usefulness and ease of use, and consequently it has a direct impact on the user's behaviour [3]. With increased diversity in the available technology, context and user's background, it may however, be relevant to understand other factors that affect adoption [14]. To identify reasons that would impact adoption, features have been divided into personal such as perceived usefulness, self-esteem and expectations and external such as social structures, regulatory environment and infrastructure, and [13]. Another model built as an extension of TAM is the psychosocial impact of assistive device scale (PIADS) [14]. The PIADS has been considered for modelling technology adoption and emphasised on personal factors and additionally considered factors such as people and society may also have an impact on usage. However both of these models have been criticised for their lack of illustration and prediction. Based on the review of these models, a comprehensive unified theory of acceptance and use of technology (UTAUT) was developed in [15]. For incorporating the user expectations relating to the system performance, the UTAUT model provides a broader definition to the perceived usefulness of the technology. Additional factors incorporated within the model include social influence and facilitating conditions. It was found that gender, age, experience and willingness of use were the main moderating influences whereas self-efficacy, attitude and anxiety did not have a direct influence on adoption. Another model built by integrating TAM with mediating factors from UTAUT is the Mobile Phone Technology Adoption Model (MOPTAM). MOPTAM was used to model personal mobile phone use by university students using features such as ease of use, usefulness, social factors, technology, and habits [16].

In recent studies it has been found that different age groups may have different perception of technology and this may affect their decisions in technology adoption [17]. With the increasing research and societal interest in understanding factors that determine adoption in older patients [18], [19], it is necessary to gain a deeper insight into technology adoption through further research. For predicting mobile phone adoption by older patients, a Senior Technology Acceptance & Adoption model for Mobile technology (STAM) extended the study on TAM for senior users [20].

There may be conflicting views between the technology developers and older users of technology based solutions. The older users generally consider that modern technology brings progress and many benefits; however, they are not sure of its benefits as they consider themselves not skilled enough to use such high-technology applications [21]. Also, they usually have a lower self-efficacy and higher technology anxiety [22]. Additionally, it has been noticed that elders do not show interest in high-technology products, however, rather value the technology that can make their daily life easier and provide added safety and security [17]. A positive impact on elders is most frequently associated with how the technology supported activities, enhanced convenience, and contained useful features [21]. While mobile phones and similar technologies may become status symbols, older users may not adopt on the sole basis that the technology is current. Moreover, the perceived usefulness of assistive technology by developers is unlikely to be considered by older users due to the associated stigma with the technology. Hence, in such contexts, it is necessary that the intended technology should address a clear need and perceived usefulness. The evaluation of awkwardness and other negative connotations were therefore included into the PIADS scale that assesses on three levels: (1) competence: the functional competence and efficiency, (2) adaptability: willingness to try out new things and take risk, and (3) self-esteem: feelings of emotional health and happiness. This 26-item scale required that a person was able to reflect and provide feedback on their perceptions, which may be deemed to be difficult for a PwD. A systematic study of the factors influencing the acceptance of electronic technologies that support aging in place by community-dwelling older adults was carried out in [23]. It was found that technology adoption was influenced by factors that can be categorised into six themes of concerns regarding technology (e.g., high cost, privacy implications and usability factors); expected benefits of technology (e.g. increased safety and perceived usefulness); need for technology (e.g. perceived need and subjective health status); alternatives to technology (e.g. help by family or spouse), social influence (e.g. influence of family, friends and professional caregivers) and characteristics of older patients (e.g., desire to age in place).

A limited amount of research has been focused on investigating technology adoption in cognitively impaired older patients [9]. Research conducted to date has focused on the issues related to general technology adoption and the relevant factors. Scherer *et al*. [14] developed a questionnaire for PwDs, aiming at identifying the most suitable assistive technology; however, the questionnaire was used to choose an assistive technology without addressing the need to predict adoption in the long term [13]. Generic factors for the use and abandonment of assistive devices for older patients were described in [5] with a focus on elderly populations with functional disabilities and impairment. Four factors were found to be significantly related to abandonment of assistive technologies in [24], namely (1) lack of consideration of user opinion; (2) easy device procurement; (3) poor device performance; and (4) changes in user needs or priorities. These findings prompted the conclusion that providing assistive technology with better service design can improve user satisfaction and reduce the rate of rejection. Factors that influence the response to healthcare robots were studied in [25] from the perspective of individual factors in addition to robotic factors in order to facilitate robotic assistance to the older patients. Nevertheless, limited attention has been directed towards the fast growing population of PwDs and their carers. One relevant study on adoption in the long term and effectiveness of telehealth was evaluated for 12 months [26]. The results of this study suggested that quality of life did not improve for chronic obstructive pulmonary disease, diabetes, or heart failure.

These findings suggest that it is necessary to conduct early evaluation of the likely adoption of these technology-based assistive solutions prior to recommending them to a PwD. Thus from the technology adoption perspective it is necessary to include features such as an individual's background like education, job and medical history obtained through diverse datasets into the adoption study. With this aim, the work described in this paper builds adoption prediction model based on the features available from the CCSMA dataset and is described next.

3. Methodology

In previous work [8], [9], we carried out preliminary studies which explored the most likely features that would affect adoption of a video-based reminder system for PwDs. The available set of features was, however, limited, as too was the analysis, particularly with regard to dependency between features. In the present work, the evaluation of the adoption models is carried out using data obtained from the CCSMA project and data gleaned from the TAUT project.

The CCSMA project involves a collaborative group of researchers at the Utah State University, Duke University Medical Center and the Johns Hopkins University, where the sample of participants was recruited in Cache County, Utah [11]. The study started in 1995 and the study enrolled 5,092 permanent residents of the county, including approximately 1033 individuals aged 60 years and older. The elderly population of Cache County has a higher educational attainment, lower incidence of chronic disease, and longer life expectancy than other related populations. Approximately every three years periodic waves of visits were conducted in the study. The study was designed to consider environmental and genetic factors related with risk for AD and other forms of dementia.

The TAUT project actively recruited subjects from the CCSMA to carry out a 12 month assessment of adoption models in association with a mobile phone based reminder application. An early stage assessment of adoption prediction from the CCSMA data was carried out as described in [27]. This research provided an initial insight into the relationship between the CCSMA feature set and the output class, *Adoption*. The work described in this paper is built upon our previous research on modelling adoption using the CCSMA data.

3.1. Previous work

A cohort of 30 people were recruited to carry out a 12 month assessment of the TAUT reminder application. For categorising individuals who were non-adopters, categorisation was performed at three different levels: (1) subjects willing to try, however, due to some reason incapable of using the reminder application are profiled with adopters using the understandings gained through the assessment process, (2) non-adopters who are categorised as capable, however, not interested in trying the app and (3) non-adopters who are incapable and uninterested. Subjects in the latter two cases were categorised using the understanding gained through questionnaires administered earlier during the recruitment process. Table 1 presents the user adoption matrix showing adopters and non-adopters.

Table 1. User adoption matrix profiling adopters and non-adopters based on capability and willingness.

<table>
<tr><th colspan="2" rowspan="2">Adoption modelling</th><th colspan="2">Capability</th></tr>
<tr><th>Yes (Proficient)</th><th>No (Non- proficient)</th></tr>
<tr><td rowspan="2">Willingness</td><td>Yes (Interested)</td><td>Adopter</td><td>Non-adopter (1)</td></tr>
<tr><td>No (Not interested)</td><td>Non-adopter (2)</td><td>Non-adopter (3)</td></tr>
</table>

The data from182 subjects was used to model adoption using 31 features extracted from the CCSMA dataset. Table 2 presents the 31 features extracted from the CCSMA dataset. Aside from age, each of these features was measured as categorical data and covered a range of areas including, gender, age, employment, MMSE score and details of a range of health conditions.

Table 2. 31 features extracted from the CCSMA dataset.

Gender	Diabetes self-endorsed
Age (Years)	Diabetes first observ.
Ethnicity	Diabetes age onset
Job category	Heart attack self-endorsed
APOE Genotype	Heart attack first observ.
APOE ε4 copy number	Heart attack age
Any variant of APOE ε4	Stroke self-endorsed
Education level	Stroke first observ.
Dementia code AD pure	Stroke Age
Dementia code Any	Hypertension self-endorsed
Last CCSMA observ.	Hypertension first observ.
CCSMA observ. date	Hypertension age onset
3MS score	High Cholesterol self-endorsed
3Ms sensory adjusted (1)	High Cholesterol first observ.
3Ms sensory adjusted (2)	High Cholesterol age onset
3Ms sensory adjusted (3)	

In contrast to the feature sets in our previous work in [8], [9], where adoption was modelled using features such as perceived utility, experience and usefulness, this work considered features related to health information. Feature selection was performed using the Information gain (IG) criterion [27]. Based on the IG criterion, features were ranked from highest to lowest and it was found that only 5 features were found to be significant: Last CCSMA observation, Any APOE ε4, APOE Genotype, Dementia code AD pure and Dementia code AD any. APOE features define the presence and type of the APOE/APOE E4 gene. Dementia codes indicates the presence of AD or other forms of dementia. A more detailed description of these features is provided later when we describe our current work. Following feature selection, classification was carried out using Naïve Bayes (NB), C4.5 decision tree (DT), and *k*NN. The results obtained indicated that considering data from the unscreened CCSMA dataset, it is possible to classify adopters and non-adopter with an F-measure of 0.79 [27].

4. Modelling adoption

In this Section we report the updated data (the study reported in [27] did not include all the participants), profile adoption and non-adoption using more relevant features based on the screening process following univariate and multivariate analysis, and considering a wider range of classification algorithms. The aim is to identify convenient features that can effectively model adoption while keeping the cost relatively low. From the updated information out of 1033 subjects in the CCSMA dataset, 346 people were screened and mailed. Following which, 21 enrolled, 146 refused, 2 were unreachable, 58 were found to be deceased, 92 could not be located, 21 were out of the area and 6 were deemed to be ineligible. Similar to the work described in the previous Section, the level of adoption is described using two classes of refuser/non-adopter and adopter, where the non-adopter class comprises both refusers and ineligible respondents, and the adopter class includes those who agreed (Table 3). The cases of codes 2, 4, 7, and 8 are not included in the study as they do not provide any information related to adoption. The 31 feature data set consisted of 173 instances comprising of 21 adopters and 152 non-adopters.

Table 3. Data dictionary used for profiling adopter and non-adopter.

	Code	Frequency	Included/Removed	Class
Not recruited	0.00	687	-	-
Enrolled in study	1.00	21	Included	Adopter
Unreachable	2.00	2	Removed	-
Refused by phone or letter	3.00	146	Included	Non-adopter
Deceased prior to study	4.00	58	Removed	-
Moved	5.00	-	-	-
Temporary moved	6.00	-	-	-
Cannot locate	7.00	92	Removed	-
Out of area	8.00	21	Removed	-
Ineligible	9.00	6	Included	Non-adopter

To understand the association between adoption and abandonment of technology, in our previous research evaluation of the mobile phone-based video streaming solution [8], [9], interviews were undertaken with various members of the research and development team [8], [9]. Through these interviews and discussions, a set of features that may impact on the adoption of the reminder app were identified. Features such as physical and cognitive ability, prior technology experience, infrastructure and perceived utility were identified as key features. Additionally, the level of encouragement given to a PwD and the role of caregiver burden were found to effect adoption. Based on these interpretations an influence diagram was suggested in [8], [9] as illustrated in Figure 1. As can be seen from Figure 1 there may be an association between the identified features and assistive technology adoption. The feature set was categorised as summary features (within the rectangles with thick lines) and independent features (within the rectangles with thin lines). The summary features may be influenced by the independent features, however, they may instead be standalone features. For example in Figure 1, the independent features previous profession and age are likely to influence technology experience (a summary feature), which consecutively may influence the perceived utility (summary feature) and subsequently affect technology adoption.

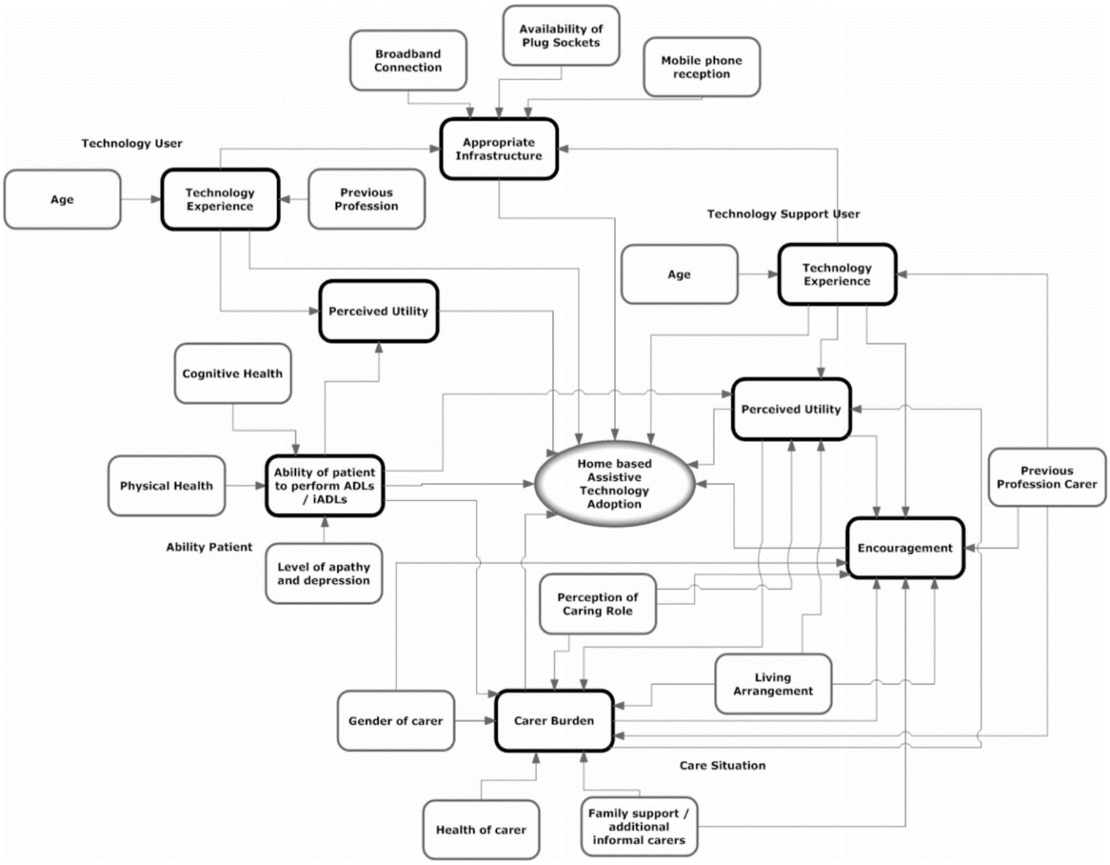

Figure 1. Influence diagram of features impacting on technology adoption [8].

Based on the CCSMA feature set, a new influence diagram has been created by including more features, which we believe can influence adoption. The new influence diagram, Figure 2, is data driven and incorporates new features that cover a range of concepts including, age, gender, MMSE score, employment and details of a number of health conditions. This influence diagram is based on the CSSMA data and incudes features available from this dataset whereas the influence diagram shown in Figure 1 is based on the view of the research team, which may not be always applicable in a real scenario. Figure 2 shows the new influence diagram created by modifying the influence diagram presented in Figure 1 and incorporating the CCSMA based feature set. For ease of comparison and to identify the progress beyond previous work, in Figure 2, the new feature set with respect to Figure 1 is shown in grey shaded rounded rectangles. In comparison to the influence diagram in Figure 1, the new influence diagram in Figure 2 has no information about the caregiver technology experience; this is because in the CCSMA feature set there is no information about the caregiver technology experience. Given the latter data were collected some years later, it was considered that the technology connectivity aspects may no longer be important and it would be beneficial to look into an individual's background for predicting adoption. Compared with the influence diagram in Figure 1, the feature, previous profession/job, is changed from an independent feature to a summary feature in the new influence diagram. The education feature is likely to impact on previous profession, however, previous profession is a feature in itself. In Figure 2, the comorbidity summary feature is summarised by independent disease features with information such as age at the time of disease onset.

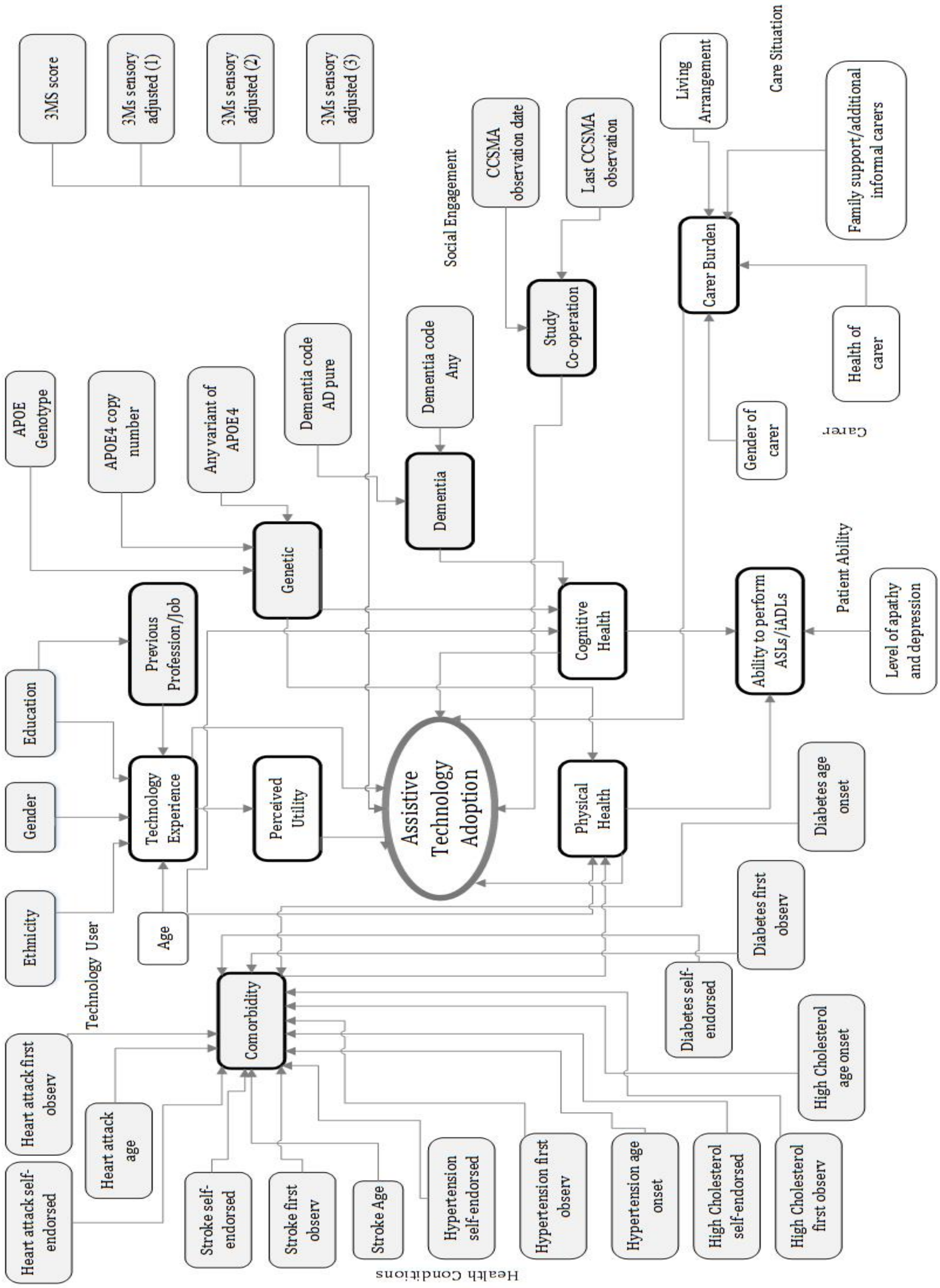

Figure 2. Influence diagram based on the CCSMA data feature set.

The features last CCSMA observed and CCSMA observed date, provide details relating to an individual's availability when the observation was taken, therefore they were included in the new influence diagram to indicate an individual's engagement with the study. It is likely that an individual who is well acquainted with the study, and associated key researchers, would be more likely to participate and adopt the technology. As indicated in [28], it is very important that the interviewers are able to convince subjects to cooperate with the survey. Based on this consideration we include a feature named study co-operation, which outlines individual willingness to partake in such studies. We believe such a parameter may be significant in an individual's decision for adoption.

4.1. Feature Selection

The CCSMA feature set, used in the previous work described in the Section 3.1, has a large set of features and it is likely that all the features may not contribute to accurate classification of the output class, *Adoption*. Therefore, it is necessary to eliminate the noisy or irrelevant features before building the model. Additionally, the classification algorithms may not scale up to the size of such a large feature set. The process of feature selection is carried out to select only those relevant features that can help in building simple and interpretable models for healthcare-based applications. Feature selection is mainly performed to eliminate irrelevant features that when included may reduce the accuracy. It is also required to look for a set of features that are relevant and cheap to collect, while giving high prediction accuracy. Based on this consideration, it is required to find a trade-off between obtaining a set of good features while still keeping the computational cost low and prediction accuracy high. The aim is to identify a subset of features that help to understand the domain better and is cheap to collect.

Feature selection is performed on the CCSMA dataset to select relevant features for the output class, *Adoption*. Features are required to be analysed at different levels: (1) features that relate directly to the adoption, and (2) features that indirectly relate to the adoption i.e. features that do not present as significant in individual pair-wise testing, however, as a group of features may significantly affect adoption due to interdependencies in these features. Initially, a pair-wise significance test is performed to find directly relevant features that would affect the output class. Out of the 31 features, the age related features such as Age, Diabetes age onset, Heart attack age, Stroke age, Hypertension age onset, and High Cholesterol age onset had continuous values while others had categorical value. In the case of small numbers, the Fisher's exact test or Chi-square [29] is used for categorical features and the Mann–Whitney-test is used for continuous data-based features. A conventional p-value = 0.05 is used for the significance thresholds. From the results obtained only the age feature with p-value = 0.029 are found to be significant. This could possibly be due to the fact that each feature has too many categorical values, leading to cells with small sample sizes. Additionally, the CCSMA feature set is large in size, which may lead to an overfitting problem and a resulting model which is excessively complex with incorrectly represented random error or noise instead of the underlying relationship. For example, in the case of Decision Trees (DT) if the dataset grows large in size while having a large number of features this would lead to the tree growing more complex. As a result, the DT built with a large number of features is perfectly suited for the existing training data, however, it may be the case that it is not able to predict unseen instances. Therefore, it is necessary to prune in order to create balance and produce parsimoneous trees with a trade-off between the number of nodes and the accuracy. Based on the new influence diagram presented in Figure 2 and the p-values obtained from the Chi-squared feature selection test, an initial manual screening undertaken to prune out irrelevant features reducing the original 31 feature set 16. These 16 features can be categorised into the broader categories of Personal, Comorbidity, Genetic, Dementia and Observation. Table 4 outlines the details of possible values for the selected 16 features which we based the adoption modelling on in the first instance.

4.1.1. Feature description

A description of the features used in the study are presented as follows.

1. *Personal features*: The personal features of an individual include gender, age, educational level and their professional background.
2. *Comorbidity features:* The comorbidity features included in the study indicate diseases such as Diabetes, Heart attack, Stroke, Hypertension and High Cholesterol. These features describe the health conditions of an individual. Either an individual never had any such conditions or they were diagnosed at some stage with these comorbidities.
3. *Genetic features:* The genetic features included in the study are APOE genotype, APOE ε4 copy number and any variant of APOE ε4.
   a. APOE genotype is a type of protein and is a key cholesterol carrier that supports lipid transport and injury repair in the brain. APOE polymorphic alleles are the main genetic determinants of AD risk: subjects having the ε4 allele are at higher risk of AD in comparison to those subjects having the more common ε3 allele, whereas the ε2 allele decreases AD risk [30].
   b. APOE ε4 copy number is called a risk-factor gene because it increases a person's risk of developing the disease. Nevertheless, inheriting an APOE ε4 allele does not mean that a person will definitely develop Alzheimer's disease. Some people with an APOE ε4 allele never get the disease, and others who develop Alzheimer's do not have any APOE ε4 alleles [31].
   c. An individual can have any variant of APOE ε4. Having a copy of the APOE ε4 gene variant doubles Alzheimer's risk for women but not for men [32].
4. *Dementia features:* Dementia code AD pure and Dementia code any indicate the presence of AD or other forms of dementia.

To have a better understanding of feature-based classification, it is necessary to have some knowledge of which features make good predictors of class membership for the classes being considered. As a benchmark for comparison the new models built with 16 features are compared with the models built with 31 features as described in Section 3.1.

Table 4. Details of the sixteen features from the CCSMA data.

Features	Details
Gender	Male = 1 Female = 2
Age (Years)	-
APOE Genotype (apoe)	22 = E2/E2 23 = E2/E3: 40% less likely 24 = E2/E4: 2.6 times more likely 33 = E3/E3: Average risk 34 = E3/E4: 3.2 times more likely 44 = E4/E4: 14.9 times more likely
APOE ε4 copy number (apoe4num)	0, 1, 2, 9 = unknown
Any variant of APOE ε4 (apoe ε4)	0 = none 1 = at least one copy
Education level (educ)	0 = No education 1 to 10 grade 11 = Eleventh grade/No diploma 12 = High School diploma or GED 13 = Some college 14 = Two years of college 15 = Three years of college 16 = College degree (B.A., B.S.) 17 = Some post-graduate work 18 = M.A., M.S. 19 = Some doctoral work 20 = Doctoral degree 97 = Refused 98 = Don't know 99 = Missing
Dementia code AD pure (padom)	1 = AD-clean 2 = AD with other dementia 3 = AD-VaD 4 = VaD without AD 5 = other dementia 9 = anycind as of x12 10 = screened normal 11 = evaluated normal 99= unable to determine
Dementia code Any (aadom)	1 =any AD 2 = VaD without AD 3 = other dementia 9 = anycind as of x12 10 = screened normal 11 = evaluated normal 99 = unable to determine
Last CCSMA observation (lastV)	1 = v1 4 = v2 7= v3 10 = v4
CCSMA observation date (lastObs)	1 = v1 2 = c1 3 = f1 4 = v2 5 = c2 6 = f2 7 = v3 8 = c3 9 = f3 10 = v4 11 = c4 12 = f4
Diabetes self-endorsed (DM)	0 = never (birth -lastVdxK) 1 = prevalent (before v1) 2 = incident (v1 to RC/Dem) 3 = post-dementia onset 4 = during PV wave(s) 9 = missing
Heart attack self-endorsed (MI)	0 = never (birth - lastVdxK) 1 = prevalent (before v1) 2 = incident (v1 to RC/Dem) 3 = post-dementia onset 4 = during PV wave(s) 9 = missing
Stroke self-endorsed (CVA)	0 = never (birth - lastVdxK) 1 = prevalent (before v1) 2 = incident (v1 to RC/Dem) 3 = post-dementia onset 4 = during PV wave(s) 9 = missing
Hypertension self-endorsed (HTN)	0 = never (birth - lastVdxK) 1 = prevalent (before v1) 2 = incident (v1 to RC/Dem) 3 = post-dementia onset 4 = during PV wave(s) 9 = missing
High Cholesterol self-endorsed (Chol)	0 = never (birth - lastVdxK) 1 = prevalent (before v1) 2 = incident (v1 to RC/Dem) 3 = post-dementia onset 4 = during PV wave(s) 9 = missing
Job category	1 = Professor, technical, manager 2 = Clerical, sales 3 = Service 4 = Agriculture 5 = Processing 6 = Machine 7 = Bench work 8 = Structural 9 = Miscellaneous 10 = Never employed

* CIND = Cognitive impairment not dementia
* VaD = Vascular Dementia
* PV wave (s) = Periodic wave of visit

4.2. Initial modelling of adoption

The 16 feature set extracted from the CCSMA database consisting of 152 non-adopters and 21 adopters is used to train and test the classification models. In the initial phase of modelling adoption, we tried to find the best fit of the candidate models. To develop the most suitable model for predicting adoption a range of popular data mining algorithms evaluated for their suitability in the prediction task [8]. Different data mining algorithms were investigated for building classification models based on their performance and fitness of the purpose. Following the parameter settings used in [8], Table 5 presents the parameter settings for the classification algorithms used to model adoption.

Table 5. Parameter settings for the classification algorithms used for building the adoption models.

Algorithm	**Parameter setting**
Neural Network (NN)	Number of hidden layers *a* = (#attributes + #classes)/2; learning rate = 0.3
C4.5 DT	Un-pruned; minimum number of leaf node = 1
Support Vector Machine (SVM)	Complexity parameter *c* = 1.0; Poly-Kernel
NB	-
Adaptive Boosting (AB)	Number of iterations = 10; base classifier: decision stump
*k*NN	*k* = 1; nearest neighbour search algorithm; linear search
Classification and Regression Trees (CART)	Heuristic search

4.2.1. Handling imbalanced classes

The CCSMA dataset consists of 152 non-adopters and 21 adopters, which is an imbalanced dataset. An imbalance of class sizes in the given data can have an impact on some of the classification algorithms, typically with a bias towards the majority class prediction [8]. Therefore in this work we also investigated the benefits gained by using resampling techniques. The given data is approximately 88% of non-adopters and 12% of adopters. Due to this imbalance in the data, the fitted models have a higher chance of incorrectly classifying most of the unknown instances to the majority class, non-adopters. Therefore for building more accurate prediction models, the imbalance in the data is addressed by applying a resampling technique.

The purpose of undertaking resampling is not to improve the accuracy and, in fact, resampling will almost always decrease the prediction accuracy [33]. Nevertheless, if we do not balance the data then the majority class will dominate which means that it will have a much greater percentage of success. It may be equally important for us to classify the minority class correctly. An extreme case of this is when there are so many examples of the majority class that the classifier assigns all cases to this class. The point of resampling is therefore to equalise the chance of the number of adopters misclassified as non-adopters and the number of non-adopters misclassified as adopters. For example, in the CCSMA data there are 152 instances of non-adopters and 21 instances of adopters; therefore if the resampling is not performed the prediction model can simply classify all the instances as non-adopters and in the extreme case it will be correct 88% of time. Thus in this case the model never classifies a patient as an adopter which defeats its main purpose. Most machine learning algorithms work best when the number of instances of each classes are roughly equal [34]. Equalising the class size allows the classification algorithms to perform the class labelling of an unknown observation solely based on the observed values of the selected features. Following class rebalance, the assignment to a specific class depends solely on the robustness of the algorithms and selected features instead the unbalanced model reflecting bias towards the majority class.

To handle this imbalance between the two classes, Synthetic Minority Over-Sampling Technique (SMOTE) is here applied. SMOTE is one of the most commonly used approaches due to its simplicity and effectiveness [35]. In this technique, new instances are created by constructing new minority class instances using a randomisation algorithm and not just by replicating the minority class. In the present work, the proportion of the data distribution is approximately 88% non-adopters and 12% adopters. The adopter minority class is given a 624% (100*(152-21)/21) boost to make it equal in size to the non-adopter class. The 16 feature set CCSMA data is resampled using the SMOTE filter in the Weka Experimenter (University of Waikato, Version 3.7.12). In Weka, the new instances of the minority class created are appended at the end of the given data. The order of appearance of the data instances are subsequently randomized to avoid overfitting. This is performed using the Randomize filter in the Weka Experimenter. Following the process of resampling, the new data is a 16 feature set with 152 instances of adopters and 152 instances of non-adopters.

4.3. Initial results

Following the resampling of the CCSMA data, the new dataset is used to build different prediction models using the classification algorithms described in Section 4.2.1. Each model is built with the 16 feature set for both the scenarios of with and without data resampling. This is performed to facilitate evaluation of the benefits gained by using resampling techniques. In this work a wider range of recognized classification algorithms is used for classification in comparison to our previous work on the CCSMA data in [27] in an effort to develop optimal classification strategies.

To investigate the relationship between classifier accuracy and the number of features, a cross validation test is performed. The goal of cross validation is to define datasets to evaluate the model in the training phase (i.e., validation datasets), in order to limit problems like overfitting, and to provide an insight into how the model will generalize to an independent dataset (i.e., an unknown dataset) [36]. Both the training and test dataset should represent samples of the underlying problem. To meet this criterion, a stratification process is useful to ensure that each output class is correctly represented in both the training and test datasets in each fold. A ten-fold cross-validation process is repeated ten times with different stratified random sampling. For evaluating the performance of each of the prediction models, the F-measure is used as a performance index. Figure 3 presents a comparative plot of averaged F-measure value of the models learned from the 16 feature set, FS_{16}, with and without SMOTE data resampling, for a range of algorithms.

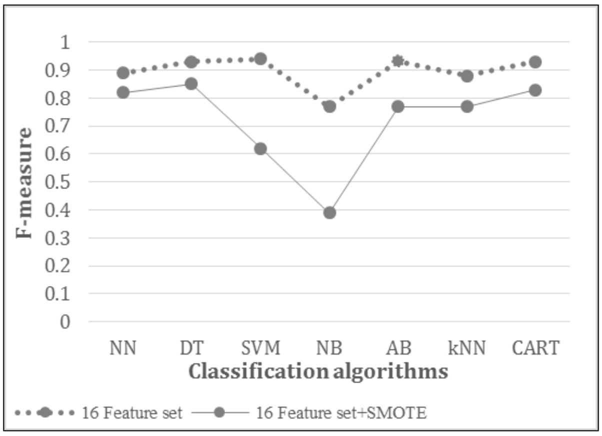

Figure 3. Averaged F-measure value of the models learned from the 16 feature set, with and without data resampling SMOTE, for a range of algorithms.

As can be viewed from Figure 3, the results for the models built with FS_{16} and SMOTE have lower accuracy in comparison to the results for FS_{16} without any resampling as discussed in the Section 4.2.1. In the previous work on the CCSMA data with the 31 feature set, FS_{31}, the F-measure index was DT = 0.79, *k*NN = 0.71, and NB = 0.42 [27]. In comparison to our previous work in [27], where the models were tested with FS_{31} and three classifiers along with non-equivalent boost of class, the current results are improved in the prediction of adopters and non-adopters. With the FS_{31}, the F-measure was DT = 0.79, kNN = 0.71, and NB = 0.42, and with FS_{16}, the F-measure index is DT = 0.85, kNN = 0.77, and NB = 0.39. Thus, it is encouraging to have fewer features with improved accuracy for further work on feature reduction. It is to be noted that the NB-based model has lower accuracy in comparison to other models. This could be due to the fact that the NB classifier struggles with the imbalanced data [37].

Additionally, we compared the prediction accuracy of FS_{31} and FS_{16} for both the cases of original and SMOTE data. In the first scenario, models are derived on the original data without handling the data imbalance for both the FS_{31} and FS_{16} feature sets. In the second scenario, the SMOTE is applied only on the training dataset and the resulting models are tested on the original data. Table 6 presents the average prediction accuracies of the models with FS_{31} and FS_{16} sets learned and tested for both the scenarios over a range of algorithms. Though the current feature set has improved results substantially, it is likely that further improvement and reduction in the number of features can still be achieved by performing further analysis on the obtained feature set.

Table 6. Average prediction accuracies of the models with FS_{31} and FS_{16} feature sets learned on original and resampled data, and tested on the original data over a range of algorithms.

	FS_{31} original data train + test (%)	**FS_{16} original data train + test (%)**	**FS_{31} SMOTE model + original test (%)**	**FS_{16} SMOTE model + original test (%)**
NN	76.3	78.61	97.69	97.10
DT	86.13	86.13	86.70	94.80
SVM	87.86	87.86	64.16	59.54
NB	41.04	69.36	40.46	36.42
AB	86.70	85.55	73.41	81.50
***k*NN**	77.46	78.61	90.17	99.42
CART	87.28	87.86	91.91	90.75

4.4. Feature categorization

After the initial screening process and reducing the feature set to 16, further feature analysis is performed to identify features that could make good predictors of class membership for the classes under investigation. Initially, the Chi-square test is performed on individual features against the output class and the models are built with our 16 feature set. As can be found from Table 4 each feature has a large number of categories, which means that the models are quite complex with a high likelihood of lower accuracy. In our initial analysis, we use Chi-squared tests for feature selection, thus reducing the feature set to 16 features (Table 4). Nevertheless, since the features typically have a large number of categories, the Chi-squared test may have few instances in some categories which weakens the test. We therefore categorise each feature into fewer values based on relevance. In Table 4, each feature is recategorised into fewer categories. For example, in the Table 4, the APOE ε4 copy number (apoe4num) feature with categorical values as 0, 1, 2, 9 = unknown is sub-categorised as NoRisk = 0 and Risk = 1, 2. A person has 0, 1, or 2 APOE ε4copy number. Having a higher APOE E4 copy number increases the risk of developing Alzheimer's disease. APOE ε4 copy number is called a risk-factor gene given that it increases a person's risk of developing the disease [31]. Table 7 presents the 16 feature set with fewer categories for each feature case. The age feature with a minimum value of 85 and a maximum value of 104 is categorised as AboveNinety and BelowOrEqualNinety. It is to be noted that the features lastV and lastObs are kept in their original form.

Table 7. 16 feature set labelled with fewer categories.

Feature	Label
gender	numeric (1, 2)
AgeLabelled	{AboveNinety, BelowOrEqualNinety}
EducationLabelled	{NoCollege, College/Higher}
JobLabelled	{KnownToTech, UnknownToTech}
ApoeLabelled	{LowerRisk, HigherRisk}
Apoe4numLabelled	{NoRisk, Risk}
APOE ε4	numeric (0, 1)
padomLabelled	{Impairment/Demntia, Normal}
aadomLabelled	{Aadom/Impairment/Dementia, Aadom/Normal}
lastV	Numeric
lastObs	Numeric
DiabetesLabelled (DM)	{None, DiagnosedAtSomeStageDM}
HeartAttackLabelled (MI)	{NoneMI, DiagnosedAtSomeStageMI}
StrokeLabelled (CVA)	{NoneCVA, DiagnosedAtSomeStagCVA}
HyperTensionLabelled (HTN)	{NoneHTN, DiagnosedAtSomeStageHTN}
HighCholestrolLabelled (CHOL)	{NoneCHOL, DiagnosedAtSomeStageCHOL}

Based on the labelling presented in Table 7, the Chi-square test is again applied and the *p*-values obtained are presented in Table 8. Here the missing information cases for a particular feature are removed from the pair-wise test as these cases are not going to add up to any information in predicting adoption. The *p*-value indicates the significance level each feature has with regard to association with the output class. A lower *p*-value indicates association between the output class and the corresponding feature. Based on this univariate test, six features are found to be significantly associated with the *Adoption* output class, namely gender, age, education, lastObs, diabetes and high cholesterol. The remaining features are not found to be significantly related to the *Adoption* output class based on the pair-wise significance test (Table 8).

Table 8. *p*-values obtained by applying the Chi-square test on the 16 features labelled data.

Feature	*p*-value
gender	0.029
AgeLabelled	0.121
EducationLabelled	0.045
JobLabelled	0.788
ApoeLabelled	0.921
Apoe4numLabelled	0.921
APOE ε4	0.865
PadomLabelled	0.914
AadomLabelled	0.914
lastV	0.734
lastObs	0.077
DiabetesLabelled (DM)	0.523
HeartAttackLabelled (MI)	0.887
StrokeLabelled (CVA)	0.415
HyperTensionLabelled (HTN)	0.827
HighCholestrolLabelled (CHOL)	0.338

4.5. Multivariate analysis

As presented in the Influence diagram (Figure 2), many features are directly and indirectly related to the output class, *Adoption*, in addition to being related to each other. This suggests that a univariate approach, such as the Chi-squared test, may not succeed in identifying all useful or relevant features. Multivariate analysis is thus carried out to determine which combination of features can work best to predict adoption, optimise the cost in relation to feature collection and reduce the computational complexity of classification. Here we use stepwise regression to perform multivariate feature selection. Stepwise regression is a greedy algorithm, widely used for feature reduction, where the algorithm selects the best feature to add, or the worst to delete, during each iteration, depending on whether it is a forward-step or backward-step approach [38]. Logistic regression is a powerful statistical way of modelling a binomial outcome (takes the value 0 or 1 such as having or not having a disease) with one or more explanatory features. It is used when the dependent variable has only two prediction classes and is categorical [39].

In our work, the 16 labelled feature set presented in Table 7 is used in a logistic regression model with conditional backward elimination, where removal is based on the probability of the likelihood-ratio statistic based on parameter estimates. Initially, the logistic regression model with conditional backward elimination is built with the original data of 173 instances with 152 non-adopters and 21 adopters. Table 9 presents a sample of a classification table obtained by running logistic regression with conditional backward elimination in the SPSS on the original data. It is to be noted that the 7 instances out of 173 had few missing values for some features, therefore, the model included only 166 instances. As can be found from Table 9 most of the instances incorrectly classified as non-adopters and the selected features are not used in the analysis. This indicates that the stepwise regression algorithm is really not working effectively and, as before, the problem of misclassification occurs due to the fact that there is an imbalance in the data. As before we use resampling to boost the minority class. Following the process described in Section 4.2.1, resampling of the data is performed and the minority class is given a 624% boost to make it equivalent to the majority class. Following this process, the logistic regression model with conditional backward elimination is fitted to the SMOTE data.

Table 9: Sample of the classification table obtained for logistic regression on the original data.

		Predicted		
		Dependent Variable		**Accuracy (%)**
	Observed	**Non-adopter**	**Adopter**	
Dependent Variable	**Non-adopter**	143	2	98.6
	Adopter	18	3	14.3
		Overall		88.0

Table 10 presents a sample of classification table obtained by running logistic regression with conditional backward elimination on the SMOTE data. As can be found from Table 10, for the SMOTE data, the feature selection process worked in that the logistic regression models using the selected features is effectively able to classify adopters and non-adopters.

Table 10: Sample of the classification table obtained for logistic regression with the SMOTE data.

	Observed	Predicted		
		Dependent Variable		Accuracy (%)
		Non-adopter	**Adopter**	
Dependent Variable	**Non-adopter**	130	22	85.5
	Adopter	26	119	82.1
		Overall		83.8

Based on the feature selection process, the set of features selected in the final logistic regression model is detailed in Table 11. In comparison to the univariate analysis where features gender, age, education, lastObs, diabetes and high cholesterol were found to be significant, multivariate analysis has added more features to the feature set, namely job, padom, lastV, heart attack, stroke and hypertension. From the results obtained using multivariate analysis it can be inferred that the professional and comorbidity features may also be useful for predicting adoption.

Table 11: Model parameters obtained by applying logistic regression.

Variable	*Model Log Likelihood*	*Change in -2 Log Likelihood*	*df*	*Sig. of the Change*
Gender	-111.982	4.097	1	.043
AgeLabelled	-114.838	9.809	1	.002
EducationLabelled	-114.177	8.487	1	.004
JobLabelled	-115.837	11.807	1	.001
PadomLabelled	-110.485	1.103	1	.294
lastV	-113.822	7.778	3	.051
lastObs	-133.939	48.011	6	.000
HeartAttackLabelled	-114.526	9.185	1	.002
StrokeLabelled	-112.848	5.830	1	.016
HyperTensionLabelled	-117.943	16.019	1	.000
HighCholestrolLabelled	-115.173	10.480	1	.001

4.6. Evaluation of Prediction models

To test the models, evaluation is performed for different scenarios. The set of classification algorithms described in Section 4.2.1 is used to build the models based on the enhanced feature set detailed in Table 11. Here we analyse two feature sets: one with all the 11 features, V_{11} (Table 11) and the other one with 9 features, V_9, which excludes lastV and lastObs. These two features are believed to be some sort of measure of social inclusion and the subject's interest in these kind of study, which may not always be easily available. The prediction accuracy is used as a performance index to evaluate the performance of each of the classifiers.

Due to an imbalance in the data, there is a chance that the model built has a bias towards for the majority class, causing a greater prediction error for the minority class. On the other hand, if the model fails to reject a possible non-adopter and classify it as adopter, in such scenarios the assistive technology is incorrectly recommended to a possible non-adopter. Successively, there will be financial implications as well as failure to effectively interact with the assistive technology can affect the mood of the PwD and hence lead to a negative impact on the quality of life of PwD and burden on the caregiver. Assessing the ease of understanding of the adoption prediction tool is significant because the end user of these tools are usually a healthcare professional. The outputs of the prediction tool need to be clinically rationalized, understood and easy to use by such healthcare professionals bearing in mind the fact they may not have a computational and technical background. The prediction models built are assessed on the overall prediction accuracy and the difference between the two types of errors (false negative and false positive classifications) to assess bias.

4.6.1. Model prediction performance

In the first scenario, models are built using the original data without handling the data imbalance. The prediction accuracies are compared between models derived using the same classification algorithm, on the given dataset with the two different features sets, V11 and V9 respectively. A 10-fold cross validation test with 10 iterations is performed to find the relationship between the classifier accuracy and features. Figure 4 presents a comparative plot of average prediction accuracies for different models with V11 and V9 feature sets learned and tested on the original data.

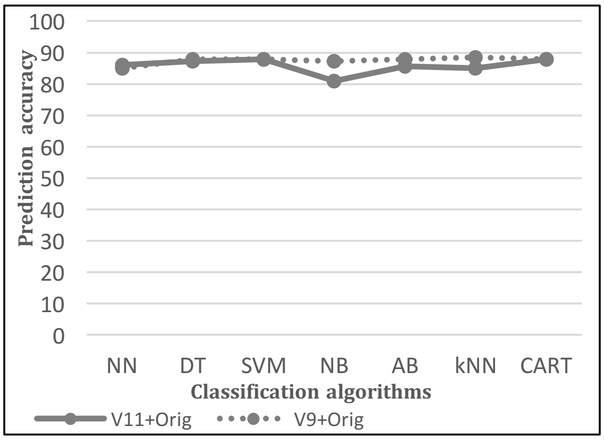

Figure 4. Average prediction accuracies of the models with V_{11} and V_9 feature sets learned and tested on the original data respectively.

Next, the models are built using the resampled data, following the resampling process described in Section 4.2.1. The resampled data consists of 152 instances of adopter and 152 instances of the non-adopter class. A 10-fold cross validation test 10 iterations is performed. Figure 5 presents a comparative plot of the average prediction accuracies for different models with V_{11} and V_9 feature sets learned and tested on the resampled data.

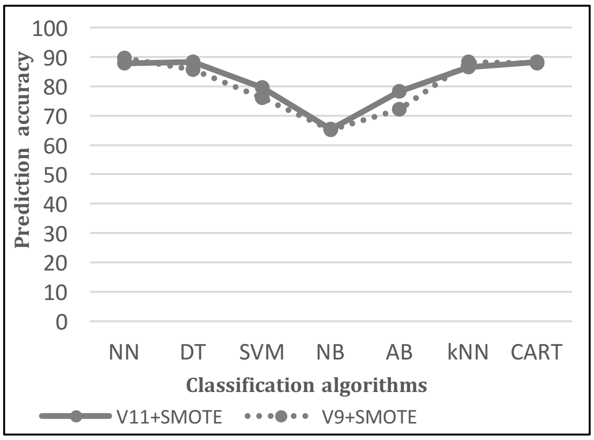

Figure 5. Average prediction accuracies of the models with V_{11} and V_9 feature sets learned and tested on the resampled data respectively.

An ideal case of testing the robustness of the built models is to train the models on resampled data and consequently test them on original data. This provides the chance to evaluate the model's performance in real world scenarios, where there may be imbalance in the observed data. The final test of the robustness of the prediction model is performed by building the models using the SMOTE data for each of the classification algorithms and obtaining the final accuracy by testing it on the original data. The model built with SMOTE data has an equal chance of classifying unseen data as a non-adopter or adopter and is actually using the feature set to classify. This is performed to obtain the final classification accuracy that we think can be achieved on an unseen dataset. The final test scenario is the SMOTE model: 304 instances (adopters = 152, non-adopters = 152) and test instances: 173 instances (adopters = 21, non-adopters = 152). Figure 6 presents a comparative plot with V_{11} and V_9 feature sets for the models built on the SMOTE data and tested on the original data.

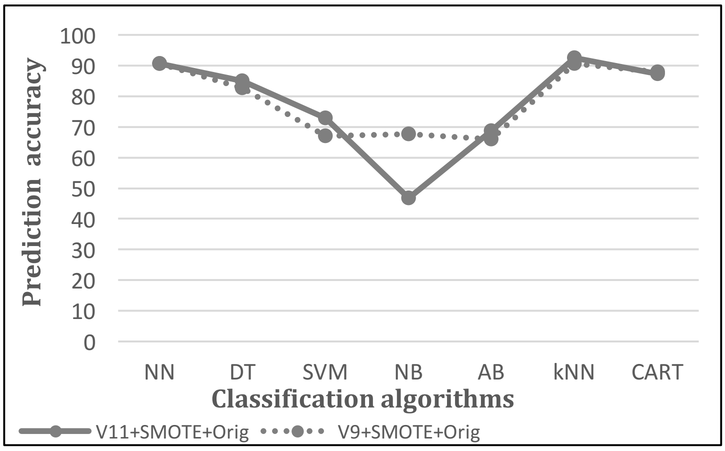

Figure 6: Average prediction accuracies of the models learned with V_{11} and V_9 feature sets trained on the SMOTE data and tested on the original data respectively.

From the classification results shown in Figures 4 and 5, the NN, DT, *k*NN and CART based prediction models have better performance in comparison to the other models across all four scenarios for both feature sets, whether or not data resampling is used to deal with the data imbalance. In the case of testing the models on unseen data, where the models are built with SMOTE data and tested on the original data *k*NN and NN based models had better performance in comparison to the other models for both the feature sets (Figure 6). Both V_{11} and V_9 feature sets have similar performance for models learned and tested on the original and resampled data (Figures 4 and 5). The same is the case when the models are built on the SMOTE data and tested on the original data. The point of analysing V_9 feature set with different classification algorithms in both cases of original and sampled data is to investigate which model performs best when the lastV and lastObs features are not available. As stated earlier these two features are considered as a measure of social inclusion and the subject's interest in these kind of study, which may not always be easily available. Figure 7 presents a pair-wise comparison between the models learned on the data with V_{11} feature set with and without data resampling.

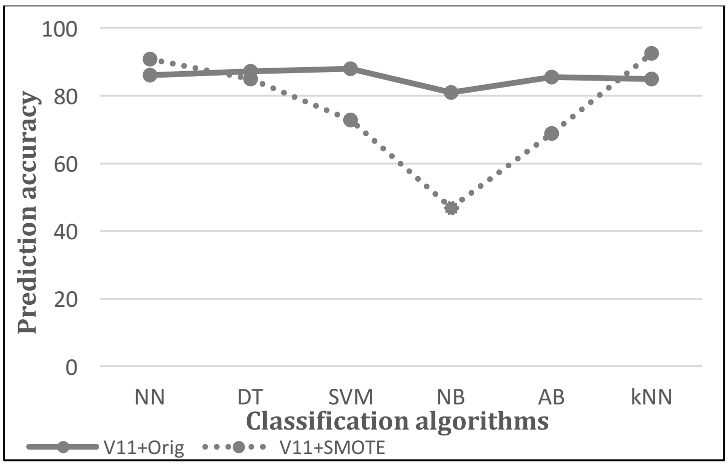

Figure 7. Averaged prediction accuracy of the models learned with V_{11} trained on data with and without data resampling and tested on the original data in both cases.

In comparison to the prediction models learned from the original dataset, the models learned using the SMOTE dataset with an adjusted probability distribution over the two output classes demonstrated a dip in prediction accuracy when tested on the original dataset, as expected. Additionally, the model built with the resampled data allows the model to have an equal chance to classify the unseen data based on the selected features in contrast to having a preference towards the majority class. Testing the SMOTE version of the built models on the original data allows to assess how the models will perform on unseen data.

4.6.2. Model prediction bias and usability

Next the model prediction bias for the two output classes in both the cases of original and resampled data is analysed. We refer to the usability of the models for the healthcare professionals who are usually non-technical. The prediction model bias to the majority class is a critical issue, particularly in scenarios where the associated cost of predicting minority is high and significantly exceeds the cost of a misclassification error for the majority class. For instance, in our case the TAUT reminder app if incorrectly suggested to a PwD can lead to a negative impact on them along with an associated financial cost. The cost estimation for misclassification is a topic for further work, requiring additional data on the cost of the financial savings and technology in addition to the evaluation of quality of life measurements. Approximating the variance between the two types of prediction errors, which is a sign of the prediction bias, is carried out. This approximation is the first step towards examining the effect of minimizing the total cost of misclassification within further work. Table 12 provides a comparison of the average prediction errors obtained between models trained on data with and without SMOTE, for V_{11} and V_9 features, respectively. As can be viewed from Table 12, in both the cases of V_{11} and V_9 feature sets, the bias towards the majority class has reduced when the resampling technique is applied.

Next the usability of the prediction models is considered. The end users of the prediction model are healthcare professionals with a non-technical background in general. Hence it is essential that the models are easy to analyse, understand and interpret. The outcomes and ease of use are the key features for these kind of healthcare-based applications. DTs are mainly valuable for healthcare-based applications as the decision making is transparent and can be visualised as trees [8]. DT-based algorithms such as CART and C4.5 DT are easily analysed and interpretable. These features allow the outcomes to be clinically rationalised and understood given the fact that the healthcare professionals have non-technical background.

Table 12. Prediction error difference between the models learned with and without SMOTE for V_{11} and V_9 features, respectively.

	Type I error	**Type II error**	**Difference**
Models with 11 features	0.0130	0.939	0.808
+ SMOTE	0.204	0.344	0.140
Models with 9 features	0.131	0.943	0.912
+ SMOTE	0.186	0.419	0.232

Conversely, the *k*NN-based models are built on the theory of finding the nearest neighbour for an unknown case on the basis of similarity measure between the unknown case and its neighbours. This characteristic of *k*NN makes its beneficial for healthcare professionals. Based on the observed values for the unknown case the outcome from the prediction model can be interpreted by the health care professional based on their past experience in the similar kind of set-up with a PwD. Contrary from the *k*NN, the output from more complex models such as NNs and SVMs is a challenging job for non-technical professionals.

5. Conclusion and future work

The question we tried to address in this work are can we accurately classify adopters and non-adopters based on the identified features, what features support or do not support prediction and predicting adoption success based on the individual's background. In this work we addressed the significance of reducing the feature set, categorisation of features into fewer sub-categories so that the models are simple and relevant for healthcare professionals to understand. The motivation for feature selection was to approximate the underlying function between the input and the output; therefore it is reasonable and important to ignore those input features with little effect on the output, so as to maintain the size of the approximator model small. The categorisation of features into fewer labels makes it easy to understand and interpret, additionally making the models simple and improving the prediction accuracy. In addition, the use of logistic regression as part of feature selection enabled us to adopt a multivariate approach that included features that were inter-related. This feature selection phase is followed by building different models using a range of classification algorithms. In addition to this the problem of data imbalance was discussed in detail along with its consequence on the prediction models. As a solution to the imbalance problem resampling is carried out to boost the adopter instances and avoid bias towards the non-adopter class. In the evaluation process the models are built on the resampled data and evaluated on the original data. This allows us to evaluate the models in advance of the unseen data taking account the imbalance within the data for the output classes. The selected features from the analysis also included the last observation and last observation date, which may not always be available. Therefore the prediction models were built with two feature sets, V_{11} and $V_{9,}$ which include features which are practically possible to collect while minimising the cost of collection. The best prediction model was obtained by using the *k*NN classification algorithm with the reduced set of features. The results obtained have improved accuracy of 92.48% over the previous prediction models built with the original 31 feature set [27]. The reduced set of labelled features with *k*NN makes the resulting model low cost, practical, and effective along with reduced complexity, which is beneficial for healthcare professionals.

In comparison to our previous work [8], where the identified set of features were gender, living arrangement, MMSE, broadband, age, mobile reception, and carer, we have more detailed feature sets that are more relevant to individual current medical conditions. The data used in this work is obtained from long-term scientific studies and represents the actual population. The data pertaining to it goes back to the same population and allows more features representing medical and scientific information. The participants in the study represent the original cohort whereas the data in the previous study was an ad hoc data collected in a basic clinical settings. In [8], we obtained very few features that might be helpful in facilitating use of technologies but not the primary factors in deciding adoption. Features such as living arrangement, broadband, and mobile reception are obvious features required for adoption but do not provide a potential insight into a PwD's ability for adoption. The work in [8] was an early stage of a study undertaken to gain a deeper insight. The current work is more reflective of a patient's decision making in long-term adoption. Decision trees were found to provide a useful alternative classification approach since they are more comprehensible by the end-user with a relatively small degradation in accuracy.

We conclude that the features identified can provide an insight into the PwD's ability to adopt technology and if the adoption is failing what could the explanation be. Features such as APOE and APOE ε4, which are genetic markers have an impact on dementia and individuals having them are likely to have reduced cognitive functionality leading to rejection of assistive technologies. The analysis did not indicate that such information is likely to be useful in terms of predicting adoption. Conversely, even though a PwD has the required infrastructure such as broadband and background he/she may not be motivated to adoption due to the stage of their dementia or other comorbidities. Hence it makes sense to include medical-based features in the adoption models. The feature sets described in the current paper provide more useful understanding into the patient ability to adopt technologies. For example, personal features such as job, education level impact on an individual's capacity to learn a new technology. Additionally, comorbidity features also have an impact on determining dementia and predicting technology adoption to some extent. Studies have found that comorbidity amongst PwDs present specific challenges for care [40]. Existence of particular comorbidities may aggravate the progression of dementia and consequently affect the user decision and their ability for adoption. Conversely, the presence of a form of dementia may badly complicate and affect the medical care of other comorbidities and is a key element of how patients' needs are estimated [40]. Including such features into the adoption model is useful in understanding a user's decision of adoption and the extent to which adoption is carried out.

Acceptance of technology is a crucial factor for successfully deploying assistive technology solutions for elderly care. Modelling user adoption from a range of parameters such as physical, environmental and social is beneficial in recommending a technology to particular users based on their background. The aim of the current work was to build models with convenient, non-invasive, and low cost features that can easily predict adoption before uptake by a potential user. With this aim we investigated features and constructed models that were intuitive and were easily understood by health care providers. Given the positive results from our previous work, the current work on the TAUT project aimed to increase the data and feature set size available for testing and training the prediction models through engagement with a larger cohort of PwDs over a longer period of time. The focus was on identifying the parameters and factors that influenced technology adoption and to develop a prediction model that may be useful to evaluate a patient-caregiver dyad and predict whether they are likely to adopt the assistive technology. These parameters can be then used as inputs for the production of more robust prediction models that could be a part of a screening process to the advantage of healthcare professionals at the point of assessing the suitability of integrating assistive technology. If they are unsuitable, a different kind of technology may be suggested or if in doubt increased PwD support may be given to overcome the identified adoption obstacles.

Future work is being undertaken to identify further differences between adopters and non-adopters by correlating usage patterns of a mobile based reminder app (e.g. number set, number missed and measures of perceived utility and usability) and including further medical history information from the UPDB dataset. Another possible topic for future work would be to combine our previous work in [8] and the current feature sets to build a more robust model that includes features such as living arrangement with medical data.

Acknowledgement

The Alzheimer's Association is acknowledged for supporting the TAUT project under the research grant ETAC-12-242841.

References

[1] C. Nugent, S. O'Neill, M. Donnelly, G. Parente, M. Beattie, S. McClean, B. Scotney, S. Mason, and D. Craig, "Evaluation of video reminding technology for persons with dementia," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, vol. 6719 LNCS, pp. 153–160.

[2] S. A. O'Neill, G. Parente, M. P. Donnelly, C. D. Nugent, M. P. Beattie, S. I. McClean, B. W. Scotney, S.

C. Mason, and D. Craig, “Assessing task compliance following mobile phone-based video reminders,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 5295–5298, 2011.

[3] D. C. Yen, C.-S. Wu, F.-F. Cheng, and Y.-W. Huang, “Determinants of users’ intention to adopt wireless technology: An empirical study by integrating TTF with TAM,” *Comput. Human Behav.*, vol. 26, no. 5, pp. 906–915, Sep. 2010.

[4] M. Chuttur, “Overview of the Technology Acceptance Model: Origins , Developments and Future Directions,” *Sprouts Work. Pap. Inf. Syst.*, vol. 9, no. 2009, pp. 1–23, 2009.

[5] L. N. Gitlin, “Why older people accept or reject assistive technology,” *Generations*, vol. 19, no. 1, pp. 41–46, 1995.

[6] R. Orpwood, C. Gibbs, T. Adlam, R. Faulkner, and D. Meegahawatte, “The design of smart homes for people with dementia - User-interface aspects,” *Univers. Access Inf. Soc.*, vol. 4, no. 2, pp. 156–164, 2005.

[7] P. Chaurasia, S. I. McClean, C. D. Nugent, and B. W. Scotney, “A duration-based online reminder system,” *Int. J. Pervasive Comput. Commun.*, vol. 10, no. 3, pp. 337–366, 2014.

[8] S. Zhang, S. I. McClean, C. D. Nugent, M. P. Donnelly, L. Galway, B. W. Scotney, and I. Cleland, “A predictive model for assistive technology adoption for people with dementia,” *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 1, pp. 375–383, 2014.

[9] S. A. O’Neill, S. I. McClean, M. D. Donnelly, C. D. Nugent, L. Galway, I. Cleland, S. Zhang, T. Young, B. W. Scotney, S. C. Mason, and D. Craig, “Development of a technology adoption and usage prediction tool for assistive technology for people with dementia,” in *Interacting with Computers*, 2014, vol. 26, no. 2, pp. 169–176.

[10] P. J. Hartin, C. D. Nugent, S. I. McClean, I. Cleland, M. C. Norton, C. Sanders, and J. T. Tschanz, “A smartphone application to evaluate technology adoption and usage in persons with dementia,” *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2014, pp. 5389–5392, 2014.

[11] J. T. Tschanz, M. C. Norton, P. P. Zandi, and C. G. Lyketsos, “The Cache County Study on Memory in Aging: Factors affecting risk of Alzheimer’s disease and its progression after onset,” *Int. Rev. Psychiatry*, vol. 25, no. 6, pp. 673–685, 2013.

[12] W. Wilkowska, S. Gaul, and M. Ziefle, “A small but significant difference - The role of gender on acceptance of medical assistive technologies,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, vol. 6389 LNCS, pp. 82–100.

[13] M. Scherer, J. Jutai, M. Fuhrer, L. Demers, and F. Deruyter, “A framework for modelling the selection of assistive technology devices (ATDs).,” *Disabil. Rehabil. Assist. Technol.*, vol. 2, no. 1, pp. 1–8, 2007.

[14] H. Day and J. Jutai, “Measuring the psychosocial impact of assistive devices: The PIADS,” *Can. J. Rehabil.*, vol. 9, no. 2, pp. 159–168, 1996.

[15] V. Venkatesh, M. G. Morris, B. D. Gordon, and F. D. Davis, “User Acceptance of Information Technology: Toward a Unified View,” *MIS Q.*, vol. 27, pp. 425–478, 2003.

[16] J. A. Van Biljon, “A Model for Representing the Motivational and Cultural Factors That Influence Mobile Phone Usage Variety,” University of South Africa (South Africa), 2006.

[17] K. Chen and A. H. S. Chan, “A review of technology acceptance by older adults,” *Gerontechnology*, vol.

10, no. 1, 2011.

[18] A. Stronge, W. Rogers, and A. Fisk, “Human factors considerations in implementing telemedicine systems to accommodate older adults,” *J. Telemed. Telecare*, vol. 13, pp. 1–3, 2007.

[19] M. Ziefle, “Age perspectives on the usefulness on e-health applications,” in *International Conference on Health Care Systems, Ergonomics, and Patient Safety (HEPS)*, 2008.

[20] K. Renaud and J. van Biljon, “Predicting Technology Acceptance and Adoption by the Elderly: A Qualitative Study,” in *Proceedings of the 2008 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries: Riding the Wave of Technology*, 2008, pp. 210–219.

[21] T. L. Mitzner, J. B. Boron, C. B. Fausset, A. E. Adams, N. Charness, S. J. Czaja, K. Dijkstra, A. D. Fisk, W. A. Rogers, and J. Sharit, “Older Adults Talk Technology: Technology Usage and Attitudes,” *Comput. Hum. Behav.*, vol. 26, no. 6, pp. 1710–1721, Nov. 2010.

[22] S. J. Czaja, N. Charness, A. D. Fisk, C. Hertzog, S. N. Nair, W. A. Rogers, and J. Sharit, “Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE).,” *Psychol. Aging*, vol. 21, no. 2, pp. 333–352, 2006.

[23] S. T. M. Peek, E. J. M. Wouters, J. van Hoof, K. G. Luijkx, H. R. Boeije, and H. J. M. Vrijhoef, “Factors influencing acceptance of technology for aging in place: A systematic review,” *International Journal of Medical Informatics*, vol. 83, no. 4. pp. 235–248, 2014.

[24] B. Phillips and H. Zhao, “Predictors of assistive technology abandonment.,” *Assist. Technol.*, vol. 5, no. 1, pp. 36–45, 1993.

[25] E. Broadbent, R. Stafford, and B. MacDonald, “Acceptance of healthcare robots for the older population: Review and future directions,” *International Journal of Social Robotics*, vol. 1, no. 4. pp. 319–330, 2009.

[26] M. Cartwright, S. P. Hirani, L. Rixon, M. Beynon, H. Doll, P. Bower, M. Bardsley, A. Steventon, M. Knapp, C. Henderson, A. Rogers, C. Sanders, R. Fitzpatrick, J. Barlow, and S. P. Newman, “Effect of telehealth on quality of life and psychological outcomes over 12 months (Whole Systems Demonstrator telehealth questionnaire study): nested study of patient reported outcomes in a pragmatic, cluster randomised controlled trial,” *BMJ*, vol. 346, 2013.

[27] I. Cleland, C. D. Nugent, S. I. McClean, P. J. Hartin, C. Sanders, M. Donnelly, S. Zhang, B. Scotney, K. Smith, M. C. Norton, and J. T. Tschanz, “Predicting technology adoption in people with dementia; initial results from the taut project,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8868, Springer Verlag, 2014, pp. 266–274.

[28] R. M. Groves and M. P. Couper, *Nonresponse in Household Interview Surveys*. Wiley, 1998.

[29] P. W. Zehna, *Probability distributions and statistics*. Allyn and Bacon, 1970.

[30] C.-C. Liu, C.-C. Liu, T. Kanekiyo, H. Xu, and G. Bu, “Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy.,” *Nat. Rev. Neurol.*, vol. 9, no. 2, pp. 106–18, 2013.

[31] ADEAR, “Alzheimer’s Disease Genetics Fact Sheet | National Institute on Aging,” *National Institutes of Health*, 2015.

[32] B. Goldman, “Having a copy of ApoE4 gene variant doubles Alzheimer’s risk for women but not for men,” *Research, Stanford News, Women’s Health, Stanford Medicine.*, 2014.

[33] P. H. Lee, “Resampling Methods Improve the Predictive Power of Modeling in Class-Imbalanced Datasets,” *Int. J. Environ. Res. Public Health*, vol. 11, no. 9, pp. 9776–9789, Sep. 2014.

[34] J.-A. Ting, A. D’Souza, S. Vijayakumar, and S. Schaal, “Efficient learning and feature selection in high-dimensional regression.,” *Neural Comput.*, vol. 22, no. 4, pp. 831–886, 2010.

[35] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” *J. Artif. Int. Res.*, vol. 16, no. 1, pp. 321–357, Jun. 2002.

[36] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, “Top 10 algorithms in data mining,” *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, 2007.

[37] N. Sobran, A. Ahmad, and Z. Ibrahim, “Classification of imbalanced dataset using conventional naive bayes classifier,” in *International Conference on Artificial Intelligence & Computer Science*, 2013, pp. 25–26.

[38] E. W. Steyerberg, M. J. C. Eijkemans, and J. D. F. Habbema, “Stepwise selection in small data sets: A simulation study of bias in logistic regression analysis,” *J. Clin. Epidemiol.*, vol. 52, no. 10, pp. 935–942, 1999.

[39] L. J. Davis and K. P. Offord, “Logistic regression.,” *J. Pers. Assess.*, vol. 68, no. 3, pp. 497–507, 1997.

[40] F. Bunn, A.-M. Burn, C. Goodman, G. Rait, S. Norton, L. Robinson, J. Schoeman, and C. Brayne, “Comorbidity and dementia: a scoping review of the literature,” *BMC Med.*, vol. 12, no. 1, pp. 1–15, 2014.