www.ijrream.in Impact Factor: 2.9463

INTERNATIONAL JOURNAL OF RESEARCH REVIEW IN ENGINEERING AND MANAGEMENT (IJRREM)

Computerized Assessment of Healthiness using Machine Learning and Contextual Approach

C. Sathish, M. Prakash, N. Rajesh, S. Prathap & K. C. Jackraj

Department of Information Technology

Er. Perumal Manimekalai College of Engineering, Hosur

ABSTRACT:

Alzheimer's disease is one of the most prevailing diseases in elderly society that leads to memory loss affecting their daily living. In this paper, an automated intelligent system is proposed to predict the multi-modal symptoms of Alzheimer's disease in order to offer appropriate actions during critical situation. To model this system machine learning techniques and contextual approach is preferred. Smart home and an intelligent system are employed to predict the symptoms of Alzheimer's disease with the help of sensors. In existing work, validation in terms of cognitive, mobility and depression states of the older adults were done using activity recognition. But the prediction of Mood plays a vital role among the multi-modal symptoms. Thus the proposed system in addition to cognitive also uses anxiety and depression states of the older adults' together helps in predicting the multi-modal symptoms. The novelty of the proposed system deals with the contextual based analysis to predict the mood using ontology approach in addition to the statistical based analysis. Using these techniques, the system measures the health assessment scores and detects a reliable change based on the assessment points in a proficient way.

Keywords - Machine learning, activity scores, mood prediction, contextual analysis, smart home, ontology approach

1. INTRODUCTION

Alzheimer's disease (AD) is one of the most prevailing diseases in developed countries that have a huge number of elderly adults getting affected. Currently more than 4 million elderly adults in India and more than 44 million elderly adults across the globe are suffering from some form of dementia including Alzheimer's disease. It has been estimated that around 115 million people will suffer from Alzheimer's disease by 2050, which can results in devastating consequences in terms of costs involved in healthcare and affects the quality of life of patients [1]. Alzheimer's disease cannot be completely cured but there are significant treatments that can delay the cognitive and behavioural symptoms as soon as they are identified and applied to the patients. Hence the early detection of disease is currently a highly significant issue as the matter of general interest. Using such methods to detect the symptoms helps in delaying and reducing the cognitive and behavioural symptoms and thereby increases the independence of the patients without the help of caregiver and reduces the health care costs.

Episodic memory impairment is the main cause of Alzheimer's disease, it also involves other symptoms related to mood, behaviour, and cognitive skills. These symptoms are measured by means of Self-report questionnaires, clinical assessments and examinations including brain imaging are conducted by doctors to measure the symptoms in all aspects. These assessments are mostly initiated when the symptoms are consistently seen for some time that could result in a delayed provision of treatment. Early identification of the disease and proper treatment on time is very important to delay and reduce the symptoms.

The proposed model works by collecting the sensor data from the smart home which is used for various factors such as estimation of sleep pattern and mobility pattern, activity recognition and to identify the duration of specific activities. The sleep pattern and mobility pattern data are used for detection of anxiety and depression. Contextual based analysis along with the anxiety detection contributes for the prediction of mood. The output of mood prediction which is the mood level data along with activity recognition and duration are given as input to the hybrid assessment model that predicts the condition of the older adult using the random forest classifier model. An assessment model is developed that hybrids various data from various sources such as activity recognition (i.e. the frequency of certain activity done by the older adult and the sequence of their activities), duration of specific activities and the prediction of mood data. The assessment model shows the result as either the older adult is healthy or at risk or experiencing difficulty.

2. RELATED WORK

Alberdi et.al. Proposed a technique that assess the possibility of detecting symptoms of Alzheimer's disease in terms of psychological, cognitive and behavioural features by using unobtrusively collected smart home behaviour data and machine learning techniques. Smart home data were collected for 2 years for about 29 older adults, the list of activities obtained from the data are given a label with the activity classes. They evaluated Mobility, cognition and mood for every six months. Regression models were used and feature selection analysis was done based on the data to predict the symptoms whereas classification models were used to identify the reliable changes in the scores predicting the symptoms. Results have showed that all mobility, cognitive, and depression symptoms can be predicted from smart home data. The limitations of this model are that it does not include additional longitudinal data that leads to insensitivity of the model and class imbalance.

P.N. Dawadi et.al proposed a technique to predict the clinical scores of smart home residents by monitoring and evaluating their daily behaviour. Clinical assessment using activity behaviour (CAAB) approach was proposed to predict the clinical scores based on their behavioural activities. Statistical features using CAAB approach is used to train a machine learning algorithm that helps in prediction of the clinical scores. Smart home sensor data were collected from 18 smart homes for two years are utilized by CAAB approach and its performance were evaluated. The limitation of this project is that it lacks in validation on huge population data that takes lots of time for computation.

J. Austin et.al. proposed a technique to assess loneliness in older adults by using wireless motion and contact sensors. Loneliness detection was done for 16 elderly adults in an unobtrusive way for 8 months of time and the accuracy of the system was measured using the regression model. The limitation of this project is that long period of time is required to assess the relationship between loneliness and behaviour to understand the effects of loneliness in a better way. Ahmad et.al. proposed a technique for detection of mild cognitive impairment (MCI) among the older adults. Smart home sensors were used unobtrusively to collect the data and machine learning techniques were used along with signal processing approach to process the data. Walking speed and other measures of the residents were calculated using the sensor for an average period of three years. Support vector machines and random forests are used to train and test the data generated from the features of different time span. The limitation of this project is that it discards lots of data that could lead to problem of over fitting.

Cheng et.al. proposed a model to predict the conversion of mild cognitive impairment (MCI) to Alzheimer's disease by classifying MCI converters from MCI non-converters. Prediction of MCI conversions to other forms of dementia are developed that uses data obtained from both forms. This method contains three main components: selection of features from a domain transfer, selection of samples from a domain transfer, and support vector machine classification. These methods were evaluated on various multiple domains from the Alzheimer's disease Support vector Machine and Kernel function are the machine learning algorithms used in this model. The limitation of this project is that it lacks in using the data more from auxiliary domains from the ADNI database that can further improves the performance.

Jie et.al. Proposed a model for longitudinal analysis of multiple different temporal data. Linear regression model was used to capture data from different time points. Data for the same region in the brain across different time-points was grouped together. An efficient optimization algorithm was developed to solve the proposed system. The proposed method improves the performance of the regression model when compared to conventional sparse learning model. Support vector Machine and Linear Regression Model were used in this project. The limitation of this paper is that it lacks in prediction based on data that requires each domain with corresponding data at each time point has constraints and minimizes the size of data for computation.

The Overall limitations studied from the review are that it lacks in comprehensive approach to detect the MCI in elderly adults and devising an online approach to automatically detect MCI is a major requirement. Clinical utility of smart home based predictions are required that helps clinicians to take informed decisions. It also lacks in finding the relationship between loneliness and behaviour data that helps in under- standing the effects of loneliness. Finally it lacks in validation on larger population and discarding too much of data resulted in problem of over fitting.

3. PROPOSED SYSTEM

The objective of this paper is to design an automated intelligent system to predict the multi-modal symptoms of Alzheimer's disease in terms of Mobility and Mood using machine learning techniques and Ontology approach. The pro- posed system of this prototype uses the artificial intelligence method and machine learning techniques and also includes the context based analysis along with the mood prediction.

The model works by collecting the sensor data from the smart home which is used for various factors such as estimation of sleep pattern and mobility pattern, activity recognition and to identify the duration of specific activities. The sleeping pattern estimation is done by calculating the total sleep disturbances of the older adult from the calculation of sleep window using the motion sensor data. Similarly the mobility pattern estimation is done by calculating the total walking times of the older adult by monitoring the sensor events across the rooms. Activity recognition is done by estimating the frequency of various activities and the sequence of all the activities by the older adult.

Finally, the duration of specific activities are calculated to find out the time spent on each activity per day. The sleep pattern and mobility pattern data are used for detection of anxiety and depression. Contextual based analysis along with the anxiety detection contributes for the prediction of mood. The output of mood prediction which is the mood level data along with activity recognition and duration are given as input to the hybrid assessment model that predicts the condition of the older adult using the random forest classifier model.

Algorithm 1: Extraction of Sleeping pattern

Input: Sensordata is a continuous sequence of event

Output: SleepPattern

Initialize globalCount, count as null Initialize e = 1

Procedure Sleep Pattern (Sensordata)

foreach Sensor events do if e“=60 then

if motionSensor.bedroom==1 then

count++

else

return count e++

Procedure Update Sleep Pattern(count)

foreach 60 minutes do

count = Sleep Pattern() globalCount = globalCount + count return global-count

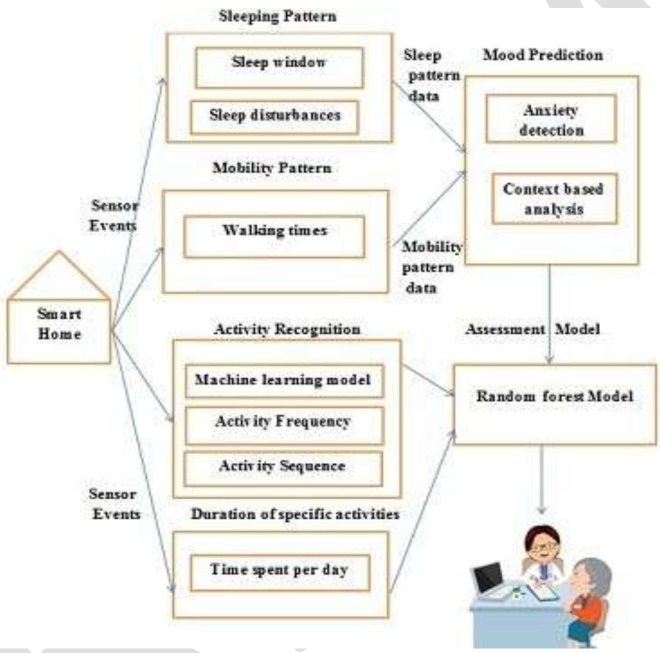


Fig. 1. Proposed Architecture for the prediction of symptoms of AD

A. Sleep Pattern Extraction

The input for this module is the Sensor data and the output is the data obtained from the sleep pattern. Heuristic approach methodology is used in this module. The module works by estimating the sleep window and determining the sleep disturbances and calculating the person's time in Bed [5]. Bedroom sensor data are monitored for every 60 minutes time interval. When the bedroom sensor gets triggered in a particular time interval (60 minutes), the count of the sleep disturbance gets incremented. The total disturbance count is estimated by adding the count of all time intervals of the complete sleep duration.

B. Mobility pattern Extraction

The input for this module is the Sensor data and the output is the data obtained from the mobility pattern. Heuristic approach methodology is used in this module. The module works by estimating walking times from sensor events at each room. Walking times of the older adults are estimated by monitoring the consecutively installed sensors that gets triggered during their movement inside a room or from one room to another. Mobility of the person that triggers the consecutive sensors within a time difference of ¡ 3 seconds are counted every 15 minutes of time. Finally the total walking times are calculated by adding all the walking times of each time interval and this total walking times shows the anxiety state of the older adult.

C. Mood Prediction

The input for this module is the data from the sleep and the mobility pattern and the output is the mood level data of the person. This module works by detecting the depression and anxiety state of the older adults using the sleep and mobility pattern data respectively [5].

Algorithm 2: Extraction of Mobility pattern

Input: Sensordata is a continuous sequence of event

Output: MobilityPattern

Initialize walking-Time, initial-Time, time-diff as null

Procedure Mobility Pattern (Sensordata)

foreach rooms k do

foreach sensors j within k do if motionSensor j==1 then

record initial-Time[j]

foreach j=1;j“n;j++ do foreach i=1;i“n;i++ do

time-diff[j][i]= initial-Time[i]-initial-Time[j]

if time-diff[j][i] “ 3 then

walking-Time++ return walking-time

Initialize total-walkingTime; Procedure Update total-walkingTime foreach 15 minutes do

walking-Time = Mobility Pattern()

total-walkingTime = total-walkingTime + walking-time

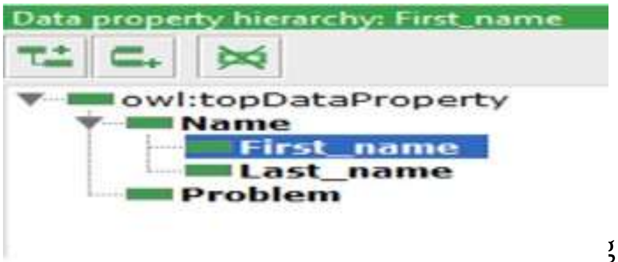
return total-walkingTime

D. Context Based Analysis

Ontology approach is the methodology used in this module. Protege 5.2.0 tool is used in this module. In this tool, classes, object properties, data properties and Individuals are defined. Classes are abstract objects used to define various entities used and a class can have any number of subclasses. Object Property is the one that relates two or more individuals of classes and data property shows the relationship between the data values. Individuals represent the various objects used in the particular domain of ontology. Here we define classes that includes Person and Emotion and an object property called "hasProblemOf" that relates the class Person with the class Emotion. We have data properties 'Name' of the older adult that includes sub property such as first /name and last /name, and the problem of the older adult. Individuals are the number of particular persons and problems that are involved in the ontology. SWRL rules are used to relate the person with the problem using a conjecture that together gives the as- sumed output. The inferred output can be seen by selecting the particular subclass that infers another subclass based on the rule defined. All the classes, properties, individuals and the relationships among them can be viewed in a graphical representation using the option called OntoGraf in Portege tool.

E. Hybrid Assessment Model

The input for this module is the training data and the output is a class per record. In this paper, the machine learning technique that has been proposed is the random forest algorithm which is a supervised learning algorithm. Random forests algorithm is an ensemble learning method works by combining the multiple decisions trees and takes random features to get more accurate prediction. Assessment dataset containing features such as the sum of the activity scores, accuracy and sequencing scores are used in this paper along with which mood scores are also taken as an additional feature and given as an input to the test data. Mood scores are obtained from the sleep and mobility pattern data and contextual analysis. Based on the input train data, the classifier predicts the output class label either as healthy or at risks or experiencing difficulty for the test data. Random forest classifier is preferred in this paper over other classifiers because of producing high accurate data value. The other advantage is that it gives a relative feature importance using which the most contributing feature can be identified. Random forest is also mainly used to handle missing values in our assessment data



g by taking the average of all the predictions retrieved from different decision trees. In the existing work also various results shows that the random forest classifier has performed well when compared to other classifiers.

Fig. 2. Classes involved in Emotion Ontology description & Fig. 3. Object Property involved in Emotion Ontology description

Figure 3 depicts the objects into properties. Object property is the one that relates the individual of person class with the individual of emotion class.

4. EXPERIMENTAL SECTION

The experimental work includes the ontology approach which has been carried out using the tool called Prote´ge´. In this tool, Emotion ontology is used to define the emotions of the person based on their problems. We have defined classes such as Person and Emotion. Based on the problems faced by the person, the Emotion of that person is inferred based on the framed rules in SWRL. The relationship between the Classes

– Person and Emotion are defined using the Object Property – “hasProblemOf”. The graphical representation of the Classes

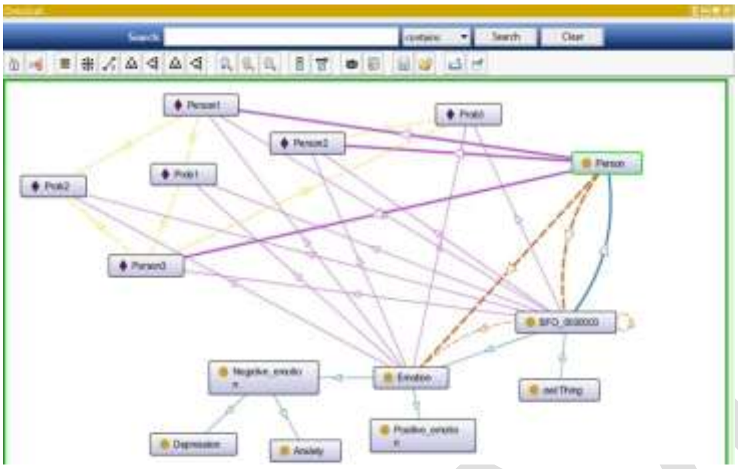


Fig. 4. Ontograph representation of various entities involved

Python 3.5 has been used to perform the machine learning classification algorithms such as Random forest and Support Vector Machine. Using the assessment dataset along with the mood scores and its corresponding class label given as a train data to both Random forest and SVM classifiers, the test data predicted results shows that the random forest classifier performs well than the SVM classifier. The accuracy, precision, recall and f1 measures produced by both classifiers.

Table I represents the comparison of results of accuracy, precision, recall and f1-score between Random forest and SVM algorithms.

Table 1: Comparison of results between random forest and Algorithms

	Accuracy	Precision	Recall	F1 score
Random Forest	0.941	0.956	0.954	0.952
SVM	0.705	0.827	0.727	0.736

5. CONCLUSION AND FUTURE WORK

The Proposed algorithm can model the hybrid assessment model by using the Random forest classifier algorithm. It measures the performance of the activities in terms of mobility and mood traits of the elderly adults. Contextual based approach is used to predict the emotions based on the problems of the elderly adults using ontology technique. The progressive estimation provides the assessment data to the doctor or caregiver so as to offer appropriate treatment based on the scores obtained in the prediction of mood of the older adults. The future work could be scaled to support a general health care system.

REFERENCES

[1] Ahmad Akl, Jasper Snoek, and Alex Mihailidis. (2015), "Unobtrusive Detection of Mild Cognitive Impairment in Older Adults Through Home Monitoring", IEEE Journal of Biomedical and Health Informatics, Vol. 21, No.2, pp. 339-348.

[2] Ahmad Akl, Babak Taati, and Alex Mihailidis. (2015), "Autonomous Unobtrusive Detection of Mild Cognitive Impairment in Older Adults", IEEE Transactions on Biomedical Engineering, Vol. 62, No. 5, pp.1383- 1394.

[3] J.Austin, H.H.Dodge, T.Riley, P.G.Jacobs, S.Thielke, and J.kaye. (2016), "A Smart-home System to Unobtrusively and Continuously Assess Loneliness in Older Adults", IEEE Journal of Translational Engineering in Health and Medicine, Vol. 4, pp. 1-11.

[4] Biao Jie, Mingxia Liu, Jun Liu, Daoqiang Zhang and Dinggang Shen. (2016), "Temporally-Constrained Group Sparse Learning for Longitu- dinal Data Analysis in Alzheimer's Disease", IEEE Transactions on Biomedical Engineering. Vol.64, No. 1, pp. 238-249.

[5] Bo Cheng, Mingxia Liu, Daoqiang Zhang, Brent C. M unsell, Dinggang Shen. (2015), "Domain Transfer Learning for MCI Conversion Predic- tion", IEEE Transactions on Biomedical Engineering, Vol.62, No. 7, pp. 1805-1817.

[6] P.N. Dawadi, D.J. Cook, and M. Schmitter-Edgecombe. (2016), "Auto- mated Cognitive Health Assessment from Smart Home-Based Behaviour Data", IEEE Journal of Biomedical and Health Informatics, Vol. 20, No.4, pp. 1188-1194.

[7] Prafulla Nath Dawadi, Diane Joyce Cook and Maureen Schmitter- Edgecombe. (2015), "Automated clinical Assessment from smart home- based behaviour data", IEEE Journal of Biomedical and Health Infor- matics, Vol. 99, No. 164, pp. 1-12.

[8] Zhou Li, Heung-II Suk, Dinggang Shen, and Lexin Li. (2016), "Sparse Multi-Response Tensor Regression for Alzheimer's Disease Study With Multivariate Clinical Assessments", IEEE Transactions on Medical Imaging, Vol. 35, No. 8, pp. 1927-1936.