ORIGINAL RESEARCH

Data-Driven Forecasting of Agitation for Persons with Dementia: A Deep Learning-Based Approach

SeyyedPooya HekmatiAthar[1] · Hilda Goins[1] · Raymond Samuel[2] · Grace Byfield[3] · Mohd Anwar[1]

Abstract

The World Health Organization estimates that approximately 10 million people are newly diagnosed with dementia each year and a global prevalence of nearly 50 million persons with dementia (PwD). The vast majority of PwD living at home receive the majority of their care from informal familial caregivers. The quality of life (QOL) of familial caregivers may be significantly impacted by their caregiving responsibilities and resultant caregiver burden. A major contributor to caregiver burden is the random occurrence of agitation in PwD and familial caregivers' lack of preparedness to manage these episodes. Caregiver burden may be reduced if it is possible to forecast impending agitation episodes. In this study, we leverage data-driven deep learning models to predict agitation episodes in PwD. We used Long Short-Term Memory (LSTM), a deep learning class of algorithms, to forecast agitations up to 30 min before actual agitation events. In particular, we managed the missing data by estimating the missing values and compensated for the class imbalance challenge by down-sampling the majority class. The simulations were based on real-world data from Alzheimer's disease (AD) caregivers and PwD dyads home environments, including ambient noise level, illumination, room temperature, atmospheric pressure (Pa), and relative humidity. Our results show the efficacy of data-driven deep learning models in predicting agitation episodes in community-dwelling AD dyads with accuracy of 98.6% and recall (sensitivity) of 84.8%.

Introduction

As the population ages, more individuals will be diagnosed with dementia. The World Health Organization (WHO) estimates that currently nearly 50 million people have dementia worldwide with approximately 10 million new cases being diagnosed annually [1]. With the progression of the disease,

✉ Mohd Anwar
  manwar@ncat.edu

[1] Department of Computer Science, North Carolina A&T State University, Greensboro, NC, USA

[2] Department of Biology, North Carolina A&T State University, Greensboro, NC, USA

[3] Department of Genetics, University of North Carolina, Chapel Hill, NC, USA

at some point persons with dementia (PwD) are unable to live alone and, thus, the PwD would require assistance in taking care of their affairs. At this stage of dementia, a caregiver is necessary. This care can be provided by an institution which caters to the elderly, or by a family member or a close friend. Informal caregiving is defined as unpaid, ongoing assistance with activities of daily living to a person with a chronic illness or disability, such as dementia [2–4]. For African Americans, the inclination is to provide care to their own loved ones in lieu of placing their family member into structured institutions [5].

PwD agitation has been confirmed as a major concern affirmed by caregivers. With agitation being the foremost issue caregivers of PwD face, it is important to explore a technological solution which can be used to distinguish between episodes of agitation and non-agitation as well as forecast agitation events using data captured prior to, during, and after an agitation event. Due to the sporadic nature of PwD agitation, it has been difficult for researchers to pinpoint when an episode may occur. However, our hypothesis

is that a link may exist between agitation and the ambient environment. In a nursing home setting which caters to dementia patients, a study by Van Hoof determined one of the causes for both restlessness and increased frequency of agitation as bright lights [6]. Furthermore, Joosse suggests that the heightened sound levels has a significant correlation to PwD agitation in a nursing home environment [7].

This research further extends the proposal by Goins et al. [8], which recommended a data-driven model based on deep learning algorithms and a combination of ambient environment data, behavioral data from both members of the dyad, and physiological data from the PwD to determine the occurrence of agitation. The research presented in this paper strives to achieve accurate results in forecasting upcoming agitation. As is often the case with health-related studies, the class imbalance issue of the dataset (e.g., significantly more instances of one class than of another) presents a challenge when applying data-driven methods. For instance, one would assume that more non-cancer screenings would result when checking for breast cancer in mammograms, fewer patients with poor arteries when assessing the risk of heart attack, or the less positive cases of COVID-19 when swabbing the general public. In this study, this problem is addressed through down-sampling the data of the majority class (i.e., non-agitation), while retaining all of the data associated with agitation events.

In this paper, a data-driven deep learning methodology is presented which, utilizing data collected within the home of a dementia dyad, distinguishes between agitation and non-agitation time segments using data collected from a single dyad of the Behavioral and Environmental Sensing and Intervention for Caregiver Empowerment (BESI) study [9–12]. Here, five environmental factors (i.e., light, audio, temperature, humidity, and air pressure) were collected and validated against the ground truth record of agitation events as reported by the caregivers. The BESI research effort consisted of three phases: (1) designed and tested the use of the tablet to collect PwD agitation information from a select few caregiver homes and verified a system for collection of environmental data from the home and behavioral data from the PwD, (2) collected and analyzed data from 12 dementia households to ascertain which ambient features were most linked to agitation and (3) provided an intervention system for which detected PwD agitation and notified the caregiver of that agitation along with carefully aligned offers of interventions to shorten the length of and decrease the severity of the detected agitation events.

Here, we implement a data-driven approach to forecast an upcoming agitation in PwD based on environmental stimuli. To this end, we use data from one of the dyads in the BESI project which was collected from 5 active relay stations for 64 days. Along with caregiver reports of PwD agitation, the data include ambient acoustic noise level, illuminance, environment temperature, atmospheric pressure, and humidity level. We preprocess the data by estimating the missing values, followed by standardization and normalization. Then, we implement and train two deep learning models without reducing dimension in the data. We also apply Principal Component Analysis (PCA) to reduce the dimensionality and train and optimize two other models based on PCA's output.

The remainder of this paper is organized as follows. The related work section provides an overview on the applications of data-driven methods in the medical domain. The data-driven modeling section discusses data-driven methods to assess caregiver burden. Then, we conduct a case study: data-driven forecasting of agitation in PwD. We conclude with a results and discussion followed by conclusion.

Related Work

Family caregivers of PwD often experience extreme stress and depression, issues of declining physical and psychological health, financial problems, and limited personal space [13–15]. In addition to declines in emotional and physical health, the colossal effect of social isolation on informal PwD caregivers is a valid concern [16]. One of the most persistent difficulties faced by caregivers, though, is the angst associated with challenging PwD behaviors [17]. Agitation is one of the major encounters that the caregiver faces and there is no set pattern that PwD agitation follows. As a result, the caregiver must be prepared at all times so that these events can be handled properly. Traditionally, agitation-related assessment and caregiver interventions are discussed with caregivers in medical settings. To the best of the authors' knowledge, there's no agitation prediction model found in the literature.

PwD agitation can be defined by three distinct characteristics: physical, verbal, and psychotic behaviors [18]. Aggressive behaviors include activities which could cause physical harm such as hitting, kicking, pushing, biting, and scratching as well as destruction of property and as well as vocal activities such as talking loudly and using inappropriate or threatening language [18, 19]. Nonaggressive behaviors include pacing, wandering and other general restlessness, as well as hiding items [17]. Physical agitation behaviors displayed by the PwD, especially those involving aggression, are the most burdensome for familial caregivers. A study on PwD patients concluded that negative PwD behavior, found in over half of the patients, not only interrupts patient care but also frustrates caregivers [20]. The literature confirms that angry and aggressive PwD behaviors are proven predictors of caregiver depression.

It has been recommended that managing PwD disturbing behaviors should begin with nonpharmacologic

interventions to improve behaviors [21]. This includes providing a safe environment, eliminating conditions in the environment that could cause stress (e.g., reduce TV sound, increase lighting, etc.) and identifying agitating or frightening situations. Interventions include redirecting the PwD, establishing appropriate sleep schedules, and developing a reward system for positive behavior [21]. Increased social support has proven to benefit the caregiver by reducing their burden [22]. Moreover, social support has been identified as a moderator of the connection between negative PwD behavior and caregiver's depressive moods; however, social support mechanisms are not consistent across all caregivers of PwD [22, 23]. Likewise, a novel method, simulated presence therapy (SPT), can be used to reduce agitation. Here, during an agitation episode, presence of a celebrated family member is simulated by playing a recording of the family member for the PwD [24].

Data-driven technology has shown promise in many areas of medical research and patient care. Moreover, due to the abundance of available data emanating from a myriad of sources, many health-related research endeavors have benefited from deep learning architectures which make use of this plethora of data. Data-driven technology has also been used in the medical environment. For instance, using electronic health records of patients with major depressive disorder (MDD), two collaborating North Carolina universities (Duke and UNC-CH) applied a data-driven approach in creating a visual representation of data from past and present patients with similar diagnosis to successfully assist physicians in decision support [25].

In Japan, where the elderly population is the highest in the world, past check-up data (waist circumference, body mass index, systolic and diastolic blood pressure, etc.), history of medications for hypertension, diabetes, and dyslipidemia were collected from pre-elderly patients (younger than 60 years old). This abundance of data along with information on previous recommendation for candidacy for a health-guidance was used to identify regular health guidance candidates. Using ensemble learning (in particular, Gradient-Boosting Decision Tree or GBDT), the results outperformed the baseline approach in the AUC by over 40%, resulting in an accuracy of 99.3% with a confidence interval of 0.993 [26].

In another study for screening and prediction of mental health, data from ubiquitous sensors, social media, and healthcare systems were fused for digital phenotyping applications. Four challenges from this data-driven approach were noted heterogeneity, volume, noise, and sparse data [27]. Depression has been studied using a data-driven approach, determining trajectory groups based on more closely spaced (i.e., weekly) severity ratings [28]. A sleep study found that the effects of insomnia may not be limited to sleep complaints, suggesting that a data-driven approach including a

list of other complaints such as quality of life, demographics, dysfunctional beliefs, childhood trauma should be considered [29]. Lastly, data-driven methodology was used as a complementary method in determining food intake of adolescents with data-driven methods showing some characteristics that were not present in the hypothesis-driven method [30]. In treatment of acute inflammation, for instance, the medicine dosage designed by data-driven methods increased survival rates from 73% to 88%. This outcome proves that data-driven models contribute to personalized treatment [31]. In another study, deep learning class of machine learning models was utilized in risk assessment and detection of diabetic foot ulcers [32].

Finally, machine learning and deep learning models have been applied successfully to medical diagnostics and screening procedures. In a lung cancer study, a convolutional neural network was trained to predict carcinoma in whole slide images. The results of the network trained on four separate datasets were within the range of 97% for the area under the curve [33]. In a similar lung cancer study, lung nodules were classified using deep learning algorithms whereas convolutional neural network yielded an accuracy of almost 90% compared to 85% for the traditional computer aided diagnostics (CADx), the deep belief network slightly exceeded CADx and the stacked denoising autoencoder only slightly trailed CADx in accuracy [34]. A similar study using computed tomography (CT) and positron emission tomography (PET) scans of 14 patients with varying stages of esophageal cancer was conducted to determine the need for surgery. Results of manually extracting the features in the traditional machine learning method and convolutional deep learning methods, which extract their own features, were compared. When traditional and deep learning techniques were used in unison 100% accuracy was achieved; whereas, an accuracy slightly less than 93% was obtained using only traditional machine learning methods [35].

Data-Driven Modeling

Environmental data provide insights for assessment of caregiver's burden. For example, some environmental events such as dimming light or increased noise levels might correspond to triggering the agitation in the PwD. Finding these events which are correlated with agitation and their impact on caregiver burden is of interest.

Monitoring the dyad's environment has been made achievable and affordable with the latest advances in digital technology. With the popularity of Internet of Things technology, a PwD is already living in a smart home or the home can feasibly be converted into a smart home. There are plenty of sensors within a smart home which actively measure different types of variables including climate (humidity

level, atmospheric pressure, temperature), illuminance, and sound levels. Alternatively, personal devices such as smart phones and tablets can be utilized. Miniaturized devices such as smart phones and wearable devices in close proximity to the dyad can be utilized for capturing biomarkers.

Assessment of Caregiver Burden

While the caregiver is first person who notices the caregiver burden, other means such as surveys, talking with family members and friends, a caregiver support group, and the medical team assigned to the PwD can be employed to ascertain the caregiver burden. However, the description of feelings that an individual provides is highly qualitative and subjective. Hence, it is difficult to develop a comprehensive framework to adequately measure caregiver burden. Data-driven methodologies can be employed to address this challenge.

Data-Driven Assessment

Different sources of data can be utilized to ascertain caregiver burden. For instance, psychometric surveys are administered to the caregiver and summarized to create a subjective synopsis of the experiences of everyday caregiving. Similarly, moments of outburst and tranquility can be ascertained using movement data which are collected by wearable devices that PwD wears. Determining the appropriate tools for caregiver burden assessment is of importance due to many possible causes of caregiver burden. These tools range from subjective caregiver input to sensors and other smart home devices.

Case Study: Data-Driven Forecasting of Agitation in PwD

Data Preprocessing

The data utilized in this paper were acquired from the BESI project, as explained in the Introduction. Three sets of environmental data were collected by each relay station placed in the home. Collected environmental data consisted of light, audio, and interior weather conditions. For this phase of the study, each of the environmental sensors gathered data throughout a time period of approximately two months. The intensity of light surrounding the sensor was measured in units of lux. For audio data, decibel levels were collected by the relay stations indicating sound intensity and not the actual voices or words from individuals. Lastly, weather data such as temperature, humidity, and barometric pressure were obtained using the relay stations. A time stamp was taken along with each data value.

The dyad chosen for this work had five relay stations dispersed throughout the home. Data were collected for a span of exactly 64 days. The raw data are provided in different segments, each of which containing a timestamp which indicates the start and duration of that segment. Each segment contains data in records which have a time offset in seconds. The summation of segment's timestamps and records' time offsets provides the absolute time of record. In BESI, the measuring devices, i.e., relay stations, were installed in different locations of the dyads house to gather a comprehensive environmental reading. The floorplan of the house in which these relay stations were installed is available; however, since the location-specific data of agitation incidents are unknown in most cases, we have averaged the data from relay stations.

An Exploratory Data Analysis (EDA) showed that (1) the sampling rate varied from segment to segment; and (2) there are some time windows with no available data. To address the inconsistent sampling rate problem, we calculated the time difference between two consecutive segments based on their timestamps; then, linearly interpolated the content of the first segment to match the duration of the time difference with the next segment. While it is possible to apply linear interpolation to address the missing data problem, the length of the missing data period might be long enough to result in inconsistency by linear interpolation. Therefore, assuming each sensory data follow the same pattern throughout the day, sensor readings from days before and after the missing value can be leveraged to estimate the missing value. The x-axis in Fig. 1 shows the days while the y-axis shows the seconds in each day. Therefore, the last record of each day is connected to the first record of the next day. Assume at time t, the sensor reading of the ith day and jth second is missing. The last sensor reading before t is a and the first sensor reading after t is b which are d_1 and d_2 seconds before and after t, respectively. The sensor readings at jth second of the days before and after, if available, are c and d. The value of t can be estimated using Eq. 1. The proof of Eq. 1 is provided in the appendix. Figure 2 shows a snapshot of temperature data before and after applying Eq. 1.

$$t = \begin{cases} \frac{d_1(b+c+d)+d_2(a+c+d)}{3(d_1+d_2)} & \text{if both } c \text{ and } d \text{ are available} \\ \frac{d_1(b+c)+d_2(a+c)}{2(d_1+d_2)} & \text{if only } c \text{ is available} \\ \frac{d_1(b+d)+d_2(a+d)}{2(d_1+d_2)} & \text{if only } d \text{ is avaibale} \\ \frac{ad_2+bd_1}{d_1+d_2} & \text{if neither } c \text{ nor } d \text{ are available} \end{cases} \quad . \quad (1)$$

After sequencing the segments and applying the interpolations, the result is a 5 by 5529,600 tensor, in which columns are representing each of five sensor readings while rows are representing 1 second timespan. Mean, Standard Deviation, Median, and Max of every 60 consecutive samples (one minute) are calculated and stored for each sensory

Fig. 1 Reference points for the estimation of missing values
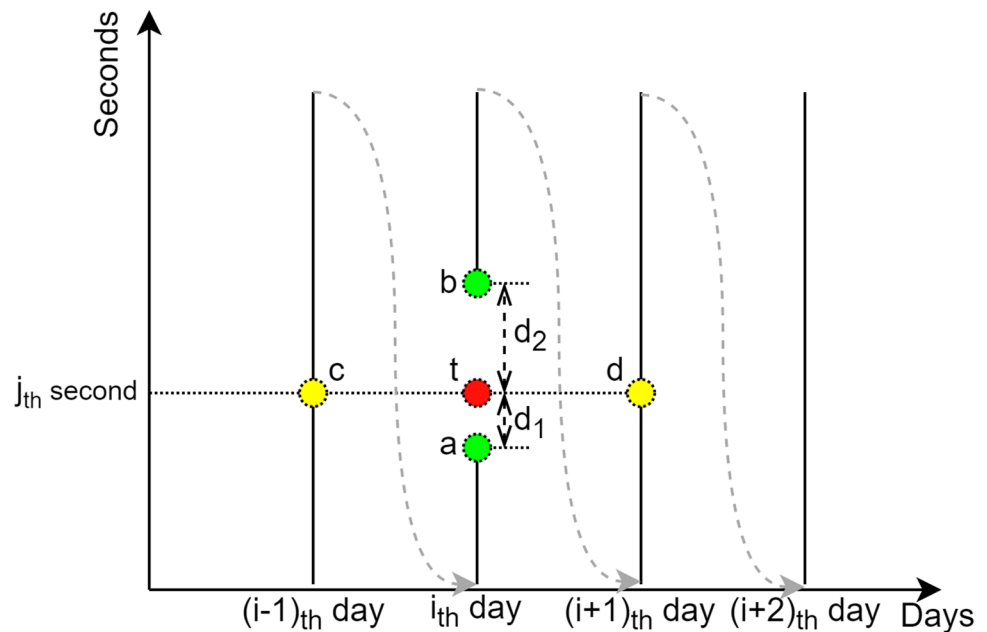
data. Finally, the values are standardized by subtracting the mean and dividing by standard deviation. The final dataset is a 20 by 88,911 tensor. After some initial models were trained, it was discovered that normalizing data, which scale the data to be in the range of [0, 1] improves the results; therefore, normalization is applied on the standardized data.

The selected dyad has recorded a total number of 41 agitations. Each agitation has a timestamp and a level which represents the severity of agitation from caregiver's perspective. We converted the classification problem to a regression problem by considering a time window with the span of one hour for each agitation (-30 min to $+30$ min). In this time window a gaussian-like function is applied. This function starts from 0 at 30 min prior to time stamp, ramps up to a peak equivalent to agitation level at the recorded time stamp and decreases back to 0 at 30 min after agitation. The combination of these 41 gaussian-like curves is considered as the ground truth for the model. The highest value caregivers reported for agitation level is 6, therefore, our values range from 0 to 6.

Considering the aforementioned one-hour time window for each agitation episode leads to 2460 non-zero samples versus 86,451 samples with the value of zero. This means 97.2% of samples belong to one class (i.e., non-agitation) while 2.8% of samples belong to other class (i.e., agitation). This is naturally a class imbalance problem. Because common practices to this issue such as collecting more data or data augmentation are not feasible here, under-sampling was applied. In this approach, each non-agitation sample is kept with a ¼ chance. This leads to having 27,965 non-agitation samples which gives us a 92% to 8% dataset, which is the lowest amount of under-sampling that achieves reasonable results.

Learning Algorithm Selection

The aim of this work is to forecast an agitation based on changes in the environmental stimuli. To this end, the models are developed to receive sensory data in the past 30 min and predict the possibility of an agitation. Since, as described in the previous section, we have 20 features, the models need to accept a 20 by 30 tensor as input and generate a scaler value which is discussed in more details in the next section. Since the data are timeseries and have temporal dependencies, Long Short-Term Memory (LSTM) is a good choice to develop this model. However, it is also possible to flatten the 20 by 30 tensor into a 600-vector tensor and develop a Multi-Layer Perceptron (MLP) model based on it. Moreover, we applied PCA to reduce the number of features from 20 to 10. This means that PCA reduced the input tensor to the size of 10 by 30.

Training Models

The LSTM model consists 32 LSTM blocks followed by a 10-neuron fully connected hidden layer which uses Rectified Linear Units (ReLU) as the activation function. The MLP model connects its 600-neuron input layer to the output via a 10-neuron hidden layer. In the MLP model, both input and hidden layers use sigmoid as the activation function. The numbers of trainable parameters are 7125 and 6131 for LSTM and MLP models, respectively. Similarly, the two other LSTM and MLP models, which are implemented
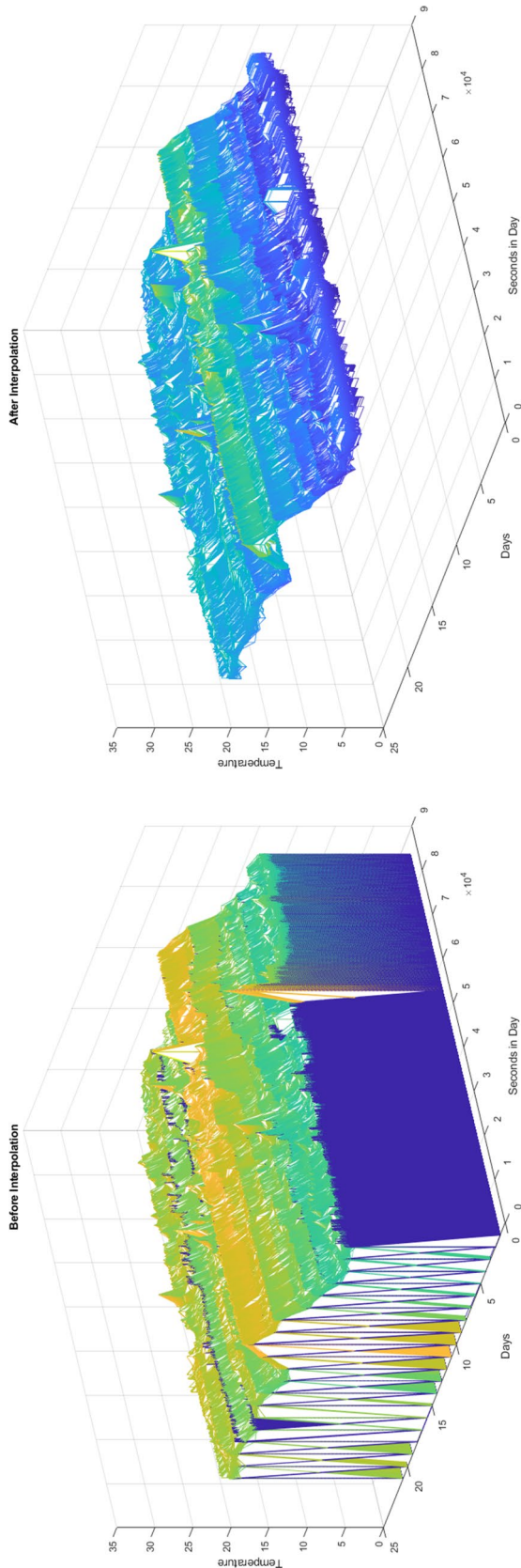
based on PCA, have 5845 and 3131 trainable parameters, respectively. Except for the input layers, the topologies of models are kept the same before and after applying PCA. Table 1 summarizes the trained models.

All models were trained using Adam optimizer and Mean-Squared Error (MSE) loss function. The 20 by 30 samples were generated from the sampled dataset, which was described in the previous section, using sliding window technique, in which, the oldest element of the sample is removed, and a new element is added. Therefore, each sample shares a subsample of 20 by 29 with the previous sample. Then, (after and before applying PCA) the samples were shuffled and 2/3 of them were selected for training and 1/3 for validation. Since nearly ¾ of the dataset was unseen by this model, the whole dataset was used as a test dataset. Figure 3 illustrates the entire pipeline for training models.

Results and Discussion

The outputs of the models were smoothed using rolling average. Since the problem was initially converted into regression, it needs to be converted back to classification. This is done by applying a threshold value of 1, i.e., if an output is equal to or greater than 1 it is considered as an agitation while all values below 1 are no agitation. The results of models are shown in the Table 2.

As explained earlier, with having more than 97% of samples in non-agitation class, the dataset is highly class imbalanced. Therefore, a model which assigns non-agitation class to all samples achieves more than 97% accuracy. For this reason, accuracy is not an informative metric here and other metrics such as precision and recall are needed. As shown in Table 2, models have achieved a wide range of precision and recall, which makes the interpretation of results difficult. Therefore, we consider the F1-Score, which is the combination of precision and recall, as the determinative performance metric. Model 4 has achieved the best F1-Score; however, we need a baseline to compare against. Assume we have four different datasets, all of which containing two classes. The first dataset is perfectly balanced; however, in the second, third, and fourth datasets, the minority class has 25%, 10%, and 1% of the population respectively. Assume the behavior of a dummy classifier is defined as:

$$\begin{cases} fn = n \times (1 - c) \\ fp = p \times (1 - c) \\ tn = n \times c \\ tp = p \times c \end{cases}, \qquad (2)$$

where c is the rate of correct classification, n and p are number of negative and positive samples respectively, $fn, fp, tn,$ and tp denote false negatives, false positives, true negatives

Fig. 2 Left: before applying Eq. 1, missing values are filled with zeros. Right: after applying Eq. 1

Table 1 Summary of trained models

	No dimension reduction		PCA	PCA
	Model 1	Model 2	Model 3	Model 4
Model topology	MLP	LSTM	MLP	LSTM
Inputs	600	20	300—reduced with PCA	10—reduced with PCA
Number of memory blocks	N/A	32	N/A	32
Neurons in hidden layers	10 (* 2 layers)	10	10 (* 2 layers)	10
Loss function	MSE	MSE	MSE	MSE
Loss value after training	0.07518997	0.11668854	0.06723165	0.06448663
Optimizer	Adam	Adam	Adam	Adam
Number of parameters	6131	7125	3131	5845

Fig. 3 The employed pipeline for training models
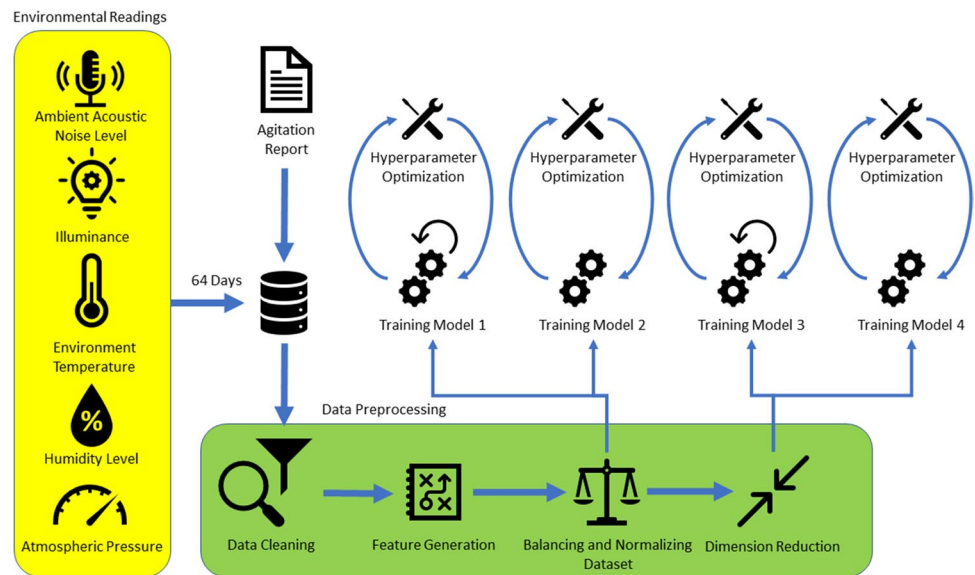
Table 2 Results of trained models

	No dimension reduction		PCA	PCA
	Model 1—MLP	Model 2—LSTM	Model 3—MLP	Model 4—LSTM
Accuracy	0.984125	0.967552	0.986319	0.98687
Precision	0.375	0.259108	0.536946	0.51581
Recall	0.219334	0.722177	0.088546	0.848091
F1-score	0.276781	0.381381	0.152022	0.641475
False negative	961	342	1122	187
False positive	450	2542	94	980
True negative	87,200	85,108	87,556	86,670
True positive	270	889	109	1044

and true positives respectively. As shown in Fig. 4, which depicts the F1-Score of this dummy classifier versus parameter c, the more unbalanced a dataset is, the lower F1-Score is for the same value of c. The horizontal lines in Fig. 4 show the F1-Score of the trained models. The intersections of these lines with datasets' curves implies how good a model is, the more right, the better. If our dataset was balanced,

the performance of model 4 could have been better as the intersection of model 4 with the balanced dataset is almost in middle range of the graph. However, for our dataset, model 4 intersects the dataset curve at the far side of the graph which means model 4 is close to the best theoretically possible model and there is little to no room for improvement. However, it is possible that the performances of the models
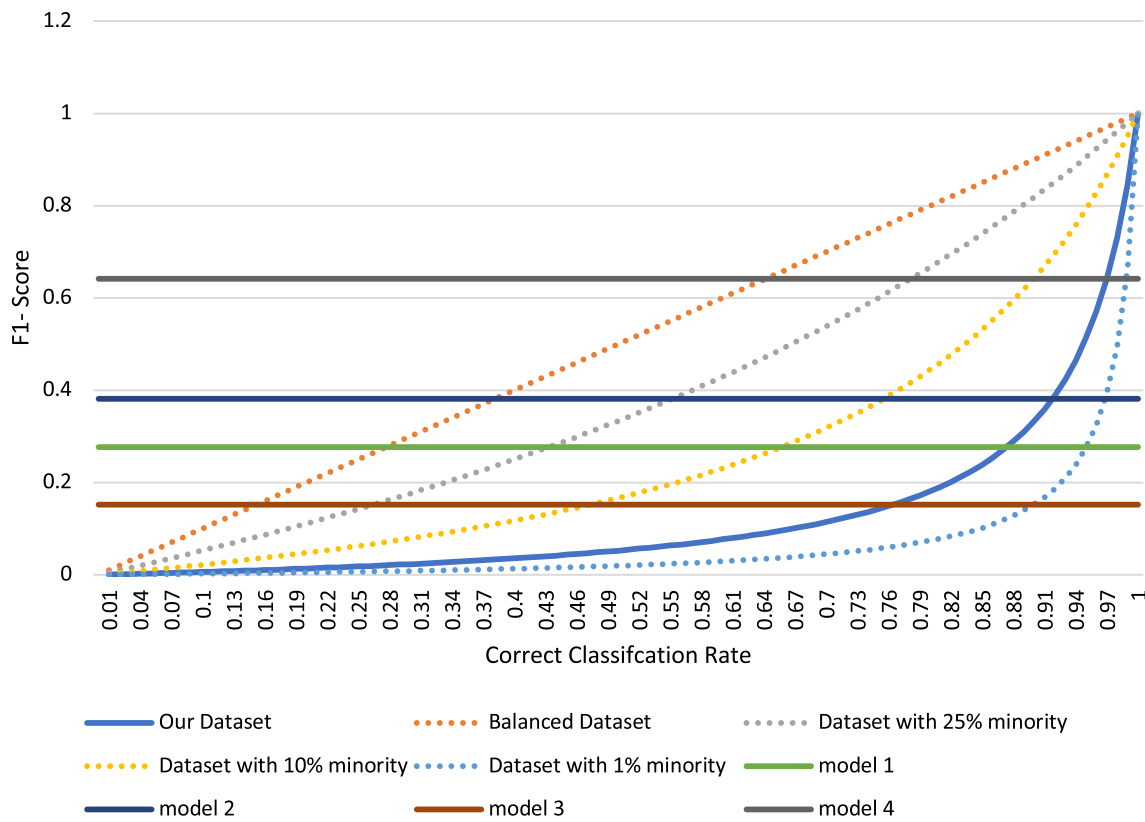
Fig. 4 Comparison of models' performances on different datasets

have reduced due to missing values and estimation used to fix it. Although we do not have any control over the missing values, a better performance in similar applications can be expected for a dataset without missing values.

Conclusion

The work we presented here can significantly reduce caregiver's burden by providing them an insight about the immediate future. Caregiver's can live their lives more peacefully knowing that they will be notified if their beloved ones are about to experience agitation. Moreover, having a model which translates environmental factors into possibility of agitation, can be used in a simulation engine to better understand how PwD react to different stimuli. This in turn provides more insight about the environmental events which are triggering agitation; thus, helping to arrange the environment with less potential for agitation and reducing caregiver burden further.

We implemented a data-driven approach to forecast an upcoming agitation in PwD based on environmental stimuli. To this end, we used data from one of the dyads in the BESI project which had five active relay stations for 64 days. The data include ambient acoustic noise level, illuminance,

environment temperature, atmospheric pressure, and humidity level. We preprocessed the data by estimating the missing values, followed by standardization and normalization. Then, we implemented and trained two Deep Learning models without reducing dimension in the data. We also applied PCA to reduce the dimensionality and implemented and trained two other models based on PCA's output. Despite some challenges such as missing values and imbalanced nature of the data, one of the trained models achieves promising results.

Appendix

In Fig. 1, the distance between t and a is d_1 units and the distance between t and b is d_2 units. Assuming the fourth case in Eq. 1 is true, where c and d are not available, estimating t using linear interpolation between a and b is:

$$t = \frac{ad_2 + bd_1}{d_1 + d_2}. \tag{3}$$

For the third cases in Eq. 1, where only d is available, assuming the importance of linear interpolation between a and b is the same is d, t can be estimated by averaging:

$$t = \frac{\frac{ad_2+bd_1}{d_1+d_2}+d}{2} = \frac{ad_2+bd_1+dd_1+dd_2}{2(d_1+d_2)} = \frac{d_1(b+d)+d_2(a+d)}{2(d_1+d_2)}.$$

$$(4)$$

Similarly, the second case in Eq. 1 is:

$$t = \frac{\frac{ad_2+bd_1}{d_1+d_2}+c}{2} = \frac{ad_2+bd_1+cd_1+cd_2}{2(d_1+d_2)} = \frac{d_1(b+c)+d_2(a+c)}{2(d_1+d_2)}.$$

$$(5)$$

For the first case in Eq. 1, where both c and d are available, the average of three values needs to be calculated:

$$t = \frac{\frac{ad_2+bd_1}{d_1+d_2}+c+d}{3} = \frac{ad_2+bd_1+c(d_1+d_2)+d(d_1+d_2)}{3(d_1+d_2)}$$

$$= \frac{d_1(b+c+d)+d_2(a+c+d)}{3(d_1+d_2)}.$$

$$(6)$$

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. World Health Organization (2020). Dementia. https://www.who.int/news-room/fact-sheets/detail/dementia. Accessed 16 Apr 2021.
2. Grabher BJ. Effects of Alzheimer disease on patients and their family. J Nucl Med Technol. 2018;46(4):335–40.
3. Wegierek AM. Taking care of a loved one who has Alzheimer's disease. Am J Alzheimers Dis Other Demen. 2012;27(7):463–7.
4. Rote SM, Angel JL, Moon H, Markides K. Caregiving across diverse populations: new evidence from the National Study of Caregiving and Hispanic EPESE. Innov Aging. 2019. https://doi.org/10.1093/geroni/igz033.
5. Moye J, Karel MJ, Gurrera RJ, Azar AR. Neuropsychological predictors of decision-making capacity over 9 months in mild-to-moderate dementia. J Gen Intern Med. 2006;21(1):78–83.
6. Van Hoof J, Aarts MP, Westerlaken AC, Schrader B, Wouters EJ, Weffers HT, Aries MB. Light therapy in smart healthcare facilities for older adults: An overview. Recent Adv Ambient Intell context-aware comput. 2015;300–307.
7. Joosse LL. Do sound levels and space contribute to agitation in nursing home residents with dementia? Res Gerontol Nurs. 2012;5(3):174–84.
8. Goins H, HekmatiAthar S, Byfield G, Samuel R, & Anwar M. Toward Data-Driven Assessment of Caregiver's Burden for Persons with Dementia using Machine Learning Models. In 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI). IEEE. 2020. p. 379–384.
9. Bankole A, Anderson M, Lach J, Smith-Jackson T, Newbold T. BESI: behavioral and environmental sensing and intervention for dementia caregiver empowerment. Am J Geriatric Psychiatry. 2019;27(3):S189–90.
10. Bankole A, Anderson MS, Homdee N, Alam R, Lofton A, Fyffe N, Goins H, Newbold T, Smith-Jackson T, Lach J. Besi: behavioral and environmental sensing and intervention for dementia caregiver empowerment-phases 1 and 2. Am J Alzheimers Dis Other Demen. 2020;35:1533317520906686.
11. Homdee N, Alam R, Hayes JA, Hamid T, Park J, Wolfe S, Goins H, Fyffe N, Newbold T, Smith-Jackson T, Bankole A, Anderson M, Lach J. Agitation monitoring and prevention system for dementia caregiver empowerment. Computer. 2019;52(11):30–9.
12. Alam R, Anderson M, Bankole A, Lach J. Inferring physical agitation in dementia using smartwatch and sequential behavior models. In 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), IEEE. 2018. p. 170–173.
13. Thompson CA, Spilsbury K, Birks Y, Barnes C, Adamson J. Systematic review of information and support interventions for caregivers of people with dementia. BMC Geriatr. 2007. https://doi.org/10.1186/1471-2318-7-18.
14. Covinsky KE, Newcorner R, Fox P. Patient and caregiver characteristics associated with depression in caregivers of patients with dementia. J Gen Intern Med. 2003;18(12):1006–14.
15. Hayslip BJ, Han GB, Anderson C. Predictors of Alzheimer's disease caregiver depression and burden: what noncaregiving adults can learn from active caregivers. Educ Gerontol. 2008;34(11):945–69.
16. Nils P. Minimizing social isolation in low-income communities: practice, and implementation strategies for dementia care. Doctoral dissertation. University of Southern California; 2019.
17. Cohen-Mansfield J. Agitated behavior in persons with dementia: the relationship between type of behavior, its frequency, and its disruptiveness. J Psychiatr Res. 2009;43(1):64–9.
18. Shahar K, Snow AL, Souchek J, Ashton CM, Kunik M. Cut-point definition of agitation. Clin Gerontol. 2004;27(1–2):15–23.
19. O'Malley KJ, Orengo CA, Kunik ME, Snow L, Molinari V. Measuring aggression in older adults: a latent variable modeling approach. Aging Ment Health. 2002;6(3):231–8.
20. He W, Sengupta M, Velkoff V, De Barros K. US census bureau, current population reports, P23–209, 65+ in the United States. US Govt. Printing Office; 2005.
21. Sadowsky CH, Galvin JE. Guidelines for the management of cognitive and behavioral problems in dementia. J Am Board Fam Med. 2012;25(3):350–66.
22. Shurgot GR, Knight BG. Influence of neuroticism, ethnicity, familism, and social support on perceived burden in dementia caregivers: pilot test of the transactional stress and social support model. J Gerontol. 2005;6(1):331–4.
23. Wang Z, Ma C, Han H, He R, Zhou L, Liang R, Yu H. Caregiver burden in Alzheimer's disease: moderation effects of social support and mediation effects of positive aspects of caregiving. Int J Geriatr Psychiatry. 2018;33(9):1198–206.
24. Zetteler J. Effectiveness of simulated presence therapy for individuals with dementia: a systematic review and meta-analysis. Aging Ment Health. 2008;12(6):779–85.
25. Mane KK, Bizon C, Owen P, Gersing K, Mostafa J, Schmitt C. Patient electronic health data-driven approach to clinical decision support. Clin Transl Sci. 2011;4(5):369–71.
26. Ichikawa D, Toki Saito T, Oyama H. Impact of predicting health-guidance candidates using massive health check-up data: a data-driven analysis. Int J Med Inf. 2017;106:32–6.
27. Liang Y, Zheng X, Zeng D. A survey on big data-driven digital phenotyping of mental health. Inf Fus. 2019;52:290–307.

28. Wardenaar KJ, Conradi HJ, Jonge P. Data-driven course trajectories in primary care patients with major depressive disorder. Depress Anxiety. 2014;31(9):778–86.

29. Benjamins JS, Migliorati F, Dekker K, Wassing R, Moens S, Blanken TF, Te Lindert BH, Mook JS, Van Someren EJ. Insomnia heterogeneity: characteristics to consider for data-driven multivariate subtyping. Sleep Med Rev. 2017;2017:36.

30. Previdelli ÁN, de Andrade SC, Fisberg RM, Marchioni DM. Using two different approaches to assess dietary patterns: hypothesis-driven and data-driven analysis. Nutrients. 2016;8(10):593.

31. Radosavljevic V, Ristovski K, Obradovic Z. A data-driven acute inflammation therapy. BMC Med Genomics. 2013;6:7.

32. Cruz-Vega I, Hernandez-Contreras D, Peregrina-Barreto H, Rangel-Magdaleno JJ, Ramirez-Cortes JM. Deep learning classification for diabetic foot thermograms. Sensors. 2020;20(6):1762.

33. Kanavati F, Toyokawa G, Momosaki S, Rambeau M, Kozuma Y, Shoji F, Yamazaki K, Takeo S, Iizuka O, Tsuneki M. Weakly-supervised learning for lung carcinoma classification using deep learning. Sci Rep. 2020;10(1):9297.

34. Sun W, Zheng B, Qian W. Automatic feature learning using multichannel Roi based on deep structured algorithms for computerized lung cancer diagnosis. Comput Biol Med. 2017;89:530–9.

35. Ferreira J, Domingues I, Sousa O, Sampaio IL, Santos JA. Classification of oesophagic early-stage cancers: deep learning versus traditional learning approaches. In: 2020 IEEE 20th international conference on bioinformatics and bioengineering (BIBE), IEEE. 2020. p. 746–751.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.