Article

Machine Learning for Predictive Modelling of Ambulance Calls

Miao Yu [1],‡ ⓘ, **Dimitrios Kollias** [2],‡ ⓘ, **James Wingate** [1] ⓘ, **Niro Siriwardena** [3] ⓘ, **Stefanos Kollias** [1] ⓘ

1   School of Computer Science, University of Lincoln, Lincoln, UK
2   School of Computing and Mathematical Sciences, University of Greenwich, London, UK
3   School of Health and Social Care, University of Lincoln, Lincoln, UK
*   Correspondence: D.Kollias@greenwich.ac.uk
‡   These authors contributed equally to this work.

Abstract: A novel machine learning approach is presented in this paper, based on extracting latent information and using it to assist decision making on ambulance attendance and conveyance to a hospital. The approach includes two steps: in the first, a forward model analyzes the clinical and, possibly, non-clinical factors (explanatory variables), predicting whether positive decisions (response variables) should be given to the ambulance call, or not; in the second, a backward model analyzes the latent variables extracted from the forward model to infer the decision making procedure. The forward model is implemented through a machine, or deep learning technique, whilst the backward model is implemented through unsupervised learning. An experimental study is presented, which illustrates the obtained results, by investigating emergency ambulance calls to people in nursing and residential care homes, over a one-year period, using an anonymized data set provided by East Midlands Ambulance Service in United Kingdom.

Keywords: predictive modelling; latent information extraction; machine learning; forward model; backward model; ambulance calls; attendance; conveyance.

1. Introduction

Most care home residents attending the Emergency Department (ED) are transported by ambulance. Ambulance conveyance to the ED is important, because it often leads to hospitalisation, and is associated with significant costs to both patient and health care system.

Many factors affect a decision to transfer a care home resident to hospital including patient, family, provider, care home, and contextual factors.

In the United States, for example, it has been estimated from a nationally representative sample that older (aged 65 years or above) nursing home residents alone accounted for 13.97 million ED visits (1.8 ED visits annually per resident) [1].

In one study, over half were not admitted to hospital and of those discharged from the ED, most had normal vital signs, nearly 1 in 5 did not have any diagnostic tests and those with injuries were 1.78 times more likely to be discharged than admitted, suggesting that these were features of preventable ED visits [1].

Over half are deemed potentially preventable using international criteria [2]. Transfers to ED are deemed potentially preventable when they "may have been avoided if optimal management of an existing condition was available at an earlier stage". Definitions usually include conditions classed as 'low acuity' (standard or non-urgent on an ED triage system such as the Manchester Triage System) and not requiring in-patient management resulting in direct discharge from ED [2].

A significant minority of patients conveyed to ED (15% in [3]) have ambulatory care sensitive conditions (ACSCs): nearly half resulted in hospital admission. Conditions with a higher risk of admission include chronic obstructive pulmonary disease, congestive heart failure, kidney/urinary tract infection, and dehydration. Various symptoms trigger transfer

to ED, the most common in one study being fatigue, lethargy or weakness, shortness of breath, and change in level of consciousness [4].

Relatives also perceive themselves to be involved in the decision to transfer as advocates for their family member [5]. Reasons for relatives encouraging conveyance including worry about nursing home care, lack of advance care planning or preparation for end of life, and poor communication or agreement between family members on goals of care [6].

Provider factors include unclear expectations, staffing capacity and capability, and limited access to multidisciplinary support or problems communicating with other decision-makers [7]. Frontline staff faced with a worsening of resident condition, insistence by family members or recommendation by a physician may deem such transfers as unavoidable, whatever subsequently happens[8] .

Some studies have shown that nursing home characteristics such as staff-resident ratio and skills as associated with transfer to ED [9]. Advance care planning and support from local health services may also reduce decisions to transfer patients to hospital [10]. Contextual factors might include time of day or day of week. Most transfers from nursing homes to EDs occur during weekdays working hours [11]. Many patients transferred to the ED are admitted (almost half in one study) and those admitted are more likely to die compared with those resident in community dwellings [12].

Our aim is to explore demographic and clinical predictors of ambulance attendance and conveyance to hospital for people residing in care homes. Preliminary results in applying machine learning to emergency medical transport datasets have been reported in [13].

In this paper we present a new methodology, composed of two steps, a forward and a backward one, in which two respective data models are generated and used to predict ambulance attendance and conveyance to hospital.

The forward step includes training of a machine, or deep learning scheme, such as a neural network, or a deep convolutional and/or recurrent network [14,15] to analyze clinical and/or demographic information and suggest conveyance, or not, of the person for which the call is made to a hospital.

The backward step is based on extraction of appropriate internal features, say features **v**, from the forward model. We then generate a model of concise representations, say **c**, through clustering of these features [15,16]. Using this model and the nearest neighbour criterion, the backward step can provide, in an efficient and transparent way, the decision on conveyance, while inferencing on the decision making process.

Recent research has focused on extracting trained network representations and using them for classification purposes, either by an auto-encoder methodology, or by monitoring neuron outputs in the convolutional or/and fully connected network layers [16,17]. Surveys about combining clustering with learning can be found in [18,19].

The novel contributions of the paper are:

i) we develop a two step approach, including a forward and a backward model, which are able to analyze data aggregated in ambulance calls in real life environments,

ii) the backward model is based on extraction of latent variables from the trained forward model, followed by appropriate clustering; thus, providing concise representations that can be analyzed in an efficient and transparent way for suggesting patient's conveyance to a hospital, or not,

iii) each concise representation set is linked to the respective input data being, therefore, able to illustrate the main (similar) cases and respective conditions on which the provided decision was based.

The rest of the paper is organized as follows. Section 2 presents the components and methods used in the proposed approach, which is also summarized in the Section. The experimental study is presented in Section 3. Discussion on the obtained results and future work are described in Section 4.

2. Materials and Methods

2.1. The Ambulance Call Data Set

Ambulance called by care home residents for transporting to the emergency department is common and involves significant costs to both the patients and the health care systems. Arguably, over half of calls may not be necessary according to some studies, which could infer significant reduction over the costs by providing an intelligent call system. There are two decisions to be made for the ambulance calls: firstly, the decision should be made if the patients would be attended by the ambulance; secondly, the decision should be made if the patients would be conveyed to hospital.

For the purpose of developing a system that assists the above-mentioned decision making, we use emergency ambulance calls to people in nursing and residential care homes over a one-year period (January-December 2018) using an anonymized data set from East Midlands Ambulance Service in United Kingdom, comprising Call and Dispatch and Clinical Records data. The data contain call categories, timings, geographical locations, clinical recordings, physiological data, treatments and outcomes of treatment, which are considered as factors informing decisions made. We have established a database for organising and manipulating the data, which is convenient to access when using machine learning tools.

This is a database using an anonymized existing data set from East Midlands Ambulance Service NHS Trust (EMAS). We do not use data taken directly from care homes, primary care or hospitals in our experimental study. We have R&D permissions from EMAS for use of the data under existing protocols for our evaluation study and a data sharing agreement between EMAS and the University of Lincoln, allowing the anonymized data set to be transferred. Data, although anonymized, are stored at the University of Lincoln on secure, suitably encrypted and password protected computer systems.

2.2. Data Pre-processing

In the database, there are clinical and non-clinical data for each patient and the decision outcomes. Data pre-processing has to be made before employing machine learning methods. It has to do data cleansing, as the data is very noisy. For example, there are large number of data instances containing missing values. In this paper we have removed instances with missing values, at least when they refer to features which are kept within the current set of experiments.

Moreover, the values of some factors are represented as text. These data have been transformed to numerical values. Data reduction was also performed, as some factors have not been deemed necessary to be included in the current analysis.

Following this procedure, we formed a data set containing around 25,500 data instances. About 20,000 of them refer to cases that were not finally transferred to a hospital, and about 5,000 of them to cases that were conveyed. Our approach has been to split the data into training (about 17,000) data, validation (about 3,400) data and test (about 5,100) data. We perform 10-fold validation on the data set, i.e., repeat the training and testing procedure, based on different splits to training, validation and testing data.

2.3. The Forward Model

In the forward model, a system is designed and delivered to guide the emergency ambulance attendance team to make a decision. The system uses the explanatory variables as input and produces the likelihood of the possible decisions, as shown in Figure 1. To develop such a system, machine learning or deep learning approaches can be employed, applied to the provided data set. The inputs include the plausible factors and the output takes a binary value, indicating, either to convey the patient(s) to hospital or not.
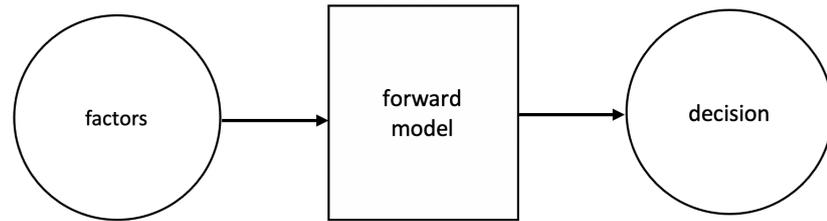
Figure 1. The forward model, for example, a DNN, is trained using the data. The inputs of the model are the clinical and/or non-clinical factors; the output is the decision whether the subject is conveyed to hospital.

2.4. The Backward Model

The backward model is based on the designed forward model; the explanatory variables which significantly informed the response are inferred from the data, as shown in Figure 2.
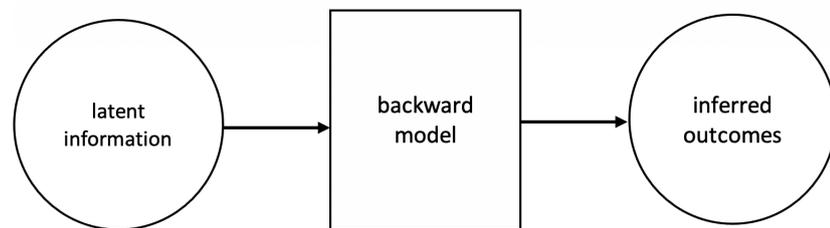

Figure 2. The backward model analyses latent information extracted from the forward model. The inputs of the model can be neurons' outputs in the trained DNN layers; the output provides insight on why the specific decision was taken.

To achieve this goal, we select to extract and further analyse the, say, M, outputs of the last fully connected layer, or last hidden layer of the trained system, e.g., artificial neural network (ANN), convolutional neural network (CNN). This is due to the fact that these outputs constitute high level, semantic extracts, based on which the trained network provides its final predictions. Other choices involve features extracted, not only from high level, but also from mid and lower level layers .

In the following we present the extraction of concise semantic information from these representations, using unsupervised learning. This information provides the explanatory variables that significantly informed the network decision.

2.5. The Proposed Approach

Ambulance called by care home residents for transportation to an emergency department is common and involves significant costs to both patients and health care systems. Arguably, over half of calls may not be necessary according to some studies, which could infer significant reduction over the costs by providing an intelligent call system. Our aim is to develop a machine learning system that assists ambulance paramedics in deciding whether there is need of conveyance or not.

Let us assume that an ANN, or DNN is trained over aggregated data set S, to predict the need of conveyance of the resident of a care home to hospital.

Let also T denote the respective test set used to evaluate the performance of the trained network. In our forward model step, we train the network using the data in S, with cardinality, say, N_s. After training, for each input k, we collect the M values of the outputs of neurons in the last fully connected layer, generating a vector $\mathbf{v}_s(k)$. A similar vector $\mathbf{v}_t(k)$ is generated when applying the trained network to each input k of the N_t test set. Thus, for all k, we get:

$$\mathcal{V}_s = \left\{ (\mathbf{v}_s(k), \ k = 1, \ldots, N_s \right\} \tag{1}$$

and

$$\mathcal{V}_t = \left\{ (\mathbf{v}_t(k),\ k = 1, \ldots, N_t \right\} \tag{2}$$

In the backward model step, we derive a concise representation of $\mathcal{V}_s$, by using a clustering procedure, based on the k-means++ algorithm [20] to generate, say, L clusters $Q = \{\mathbf{q}_1, \ldots, \mathbf{q}_L\}$ through minimization of the following function:

$$\widehat{Q}_{k\text{-means}} = \arg \min_Q \sum_{i=1}^{L} \sum_{\mathbf{v}_s \in V_s} ||\mathbf{v}_s - \mu_i||^2 \tag{3}$$

in which μ_i denotes the mean of $\mathbf{v}$ values belonging to cluster i. For each cluster i, we then compute the corresponding cluster center $\mathbf{c}(i)$, thus defining the set of cluster centers C, generating a concise representation and conveyance prediction model.

$$\mathcal{C} = \left\{ (\mathbf{c}(i),\ i = 1, \ldots, L \right\} \tag{4}$$

The procedure, of using data set S to generate the set of cluster centers C is illustrated in Figure 3, through clustering of the latent variables extracted from the, e.g., DNN system.

The derived representation will generally consist of a small number of cluster centers, for which we can examine the respective input experimental variables.

Let us now focus on using the set C for predicting the conveyance status in new subject cases, e.g., those included in the test data set T. For each input in T, we compute the $\mathbf{v}_s$ value. We then calculate the euclidean distance of this value from each cluster center in C and classify the test input to the category of the closest cluster center. As a result, we classify each test input to a respective category, thus predicting the subject's conveyance status.
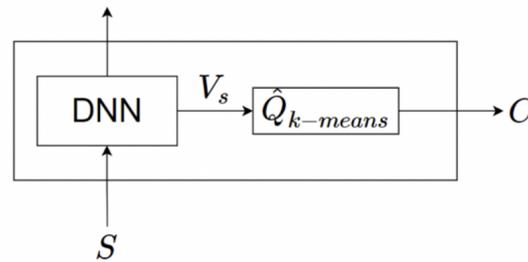

Figure 3. Extraction of latent variables $\mathbf{V_s}$ from the trained ANN/DNN (forward model) and generation of the representation set C (backward model).

It should be mentioned that, using this approach, we can predict a new subject's status in a rather efficient and transparent way. At first, only L distances between M-dimensional vectors have to be computed and the minimum of them be selected. Then, it can be inferred why the specific conveyance decision was taken, through correlation with the explanation variables corresponding to the selected cluster center.

3. Experimental Results

3.1. The final data set

The experimental study was based on the ambulance call data set which was described in the former Section. In particular, we focused our research on eight explanatory variables, which are listed in Table 1, including subject's consciousness, heart rate, systolic blood pressure, respiratory rate, subject's (destination) condition, oxygen saturation (SPO2), glucose and diastolic blood pressure.

Table 1. Correspondence of explanatory variables to subject's medical condition.

Explanatory Variables	Definitions
1	Consciousness
2	Heart Rate
3	BP Systolic
4	Respiratory Rate
5	Destination Condition
6	SP02
7	Glucose
8	BP Diastolic

Focusing on these variables, we selected a large part of the data set, consisting of about 3,300 cases in which the subject was conveyed to hospital (Category 0) and 21,300 cases in which no conveyance was made (Category 1). We split the data so that 2,034 and 13,474 cases were selected as training data, respectively, in each category. Accordingly, 700 and 4,200 cases were selected as test data, with the rest being used for validation purposes.

As has been mentioned, in the forward model step we used 10-fold validation, computing average values for the selected performance measures, which were accuracy and F1 score (per class and total) [14]. The values of all variables were normalized in the interval [0, 1] before performing the modelling steps.

3.2. Developing the forward model

Different machine learning techniques can be used in the forward model estimation step, where a system is trained to predict the category to which the input data correspond. We trained three different classifiers, i.e., Random Forest Regression (RFR), Support Vector Machines (SVM) and Artificial Neural Networks, as well as convolutional neural networks, to determine whether an ambulance needs to be called out or not based on each subject's information.

Since the provided data set was unbalanced, in all experiments we performed data augmentation in category 0, for example, by using same number of samples from the two classes in each batch of the learning algorithms.

RFR belongs to the category of ensemble learning algorithms [21]. As a base learner of the ensemble, RFR uses decision trees. RFR constructs independent regression trees, with a bootstrap sample of the training data being chosen at each regression tree. As a consequence, the regression tree continuously grows until it reaches the largest possible size. Final prediction is a weighted average of all regression tree predictions.

SVM arises from a nonlinear generalization of the Generalized Portrait algorithm [22]. SVM uses a kernel function, which allows to project the original data into a higher dimensional space to be linearly separable. The radial basis function is an efficient kernel choice, which has been used in our work.

Multilayer perceptrons (MLP) have been the main Artificial Neural Network (ANN) architecture used for supervised learning and classification tasks in the past [23]; they consist of multiple fully connected layers of neurons, with feedforward spread of information. Their training is performed with the backpropagation algorithm. In our work, the MLP was a two-layer fully connected neural network, with the first layer containing 16, or 50, neurons and the second layer containing 16 neurons, all with 'Relu' activation functions, followed by the output layer composed of two neurons corresponding to the two categories. Other network architectures have also been tried, but with no obvious performance improvement.

Convolutional neural networks (CNN) have been the main Deep Neural Network (DNN) architecture in the last decade [14]. In this work we have used networks consisting of multiple convolutional layers, followed by fully connected layers. Some choices have been to use two convolutional layers, each containing 64 activation maps, followed by either a fully connected layer containing 16 neurons and/or the output layer.

We have trained the above architectures with the data set described above. The obtained prediction accuracy in all cases is shown in Table 2. It can be seen that the CNN model has a superior performance and has been selected as forward model in our approach.

Table 2. Different methodologies along with their performance).

Methods	Accuracy of Conveyance Category	Accuracy of Non-Conveyance Category	Total Accuracy	Weighted F1 Score
RFR	0.619	0.773	0.749	0.731
SVM	0.62	0.774	0.75	0.732
ANN	0.606	0.794	0.765	0.748
CNN	0.668	0.819	0.801	0.783

3.3. Developing the backward model

We followed the procedure described in the former section and extracted the vector with elements the 16 neuron outputs of the CNN fully connected layer, to be used as latent information in the backward model. We then applied the clustering procedure on the training data set and used the validation/test data set to select the optimal number of clusters.

Generation of 20 clusters, 10 per category, was the outcome of this part of our approach. We then computed the respective cluster centers, each being represented by a vector in a 16-dimensional space. The backward model is composed of these 20 vectors, together with the respective 8-dimensional explanatory variable input vector. Tables 3 and 4 show the 8 explanatory variables for each cluster center respectively for the conveyance and non-conveyance categories.

Table 3. Conveyance category (respective inputs for the 10 cluster centres).

Cluster Number	Explanatory Variables							
0	0.0	0.237	0.327	0.072	0.2	0.97	0.053	0.472
1	0.0	0.26	0.301	0.125	0.2	0.98	0.042	0.372
2	1.0	0.0	0.0	0.0	1.0	0.0	0.038	0.0
3	0.0	0.27	0.291	0.080	0.2	0.96	0.0356	0.377
4	0.0	0.297	0.252	0.080	0.6	0.98	0.041	0.352
5	0.0	0.283	0.289	0.089	0.2	0.92	0.0648	0.292
6	0.0	0.247	0.286	0.080	0.2	0.98	0.059	0.292
7	1.0	0.267	0.267	0.063	0.2	0.99	0.038	0.362
8	0.0	0.277	0.37	0.741	0.2	0.98	0.044	0.417
9	0.5	0.363	0.169	0.067	1.0	0.97	0.111	0.261

Table 4. Non-conveyance category (respective inputs for 10 cluster centres).

Cluster Number	Explanatory Variables							
0	1.0	0.28	0.329	0.071	0.2	0.98	0.047	0.402
1	0.0	0.313	0.34	0.080	0.0	0.98	0.058	0.477
2	0.0	0.417	0.291	0.143	0.0	0.93	0.045	0.392
3	1.0	0.253	0.293	0.071	0.0	0.99	0.038	0.412
4	0.5	0.253	0.284	0.071	0.0	0.94	0.066	0.382
5	0.5	0.387	0.383	0.107	0.2	0.97	0.051	0.452
6	0.0	0.233	0.286	0.071	0.0	0.99	0.054	0.302
7	0.5	0.317	0.226	0.089	0.2	0.96	0.029	0.302
8	0.0	0.267	0.209	0.089	0.0	0.97	0.051	0.362
9	1.0	0.31	0.234	0.089	0.0	1.0	0.062	0.372

To illustrate how the training input data were split in these clusters, Tables 5 and 6 show the number of training data in each of the 10 clusters, respectively for each category.

Table 5. Conveyance clusters: number of elements.

Cluster Number	Number of Elements
0	167
1	407
2	7
3	660
4	37
5	121
6	616
7	5
8	10
9	4

Table 6. Non-conveyance clusters: no of elements.

Cluster Number	Number of Elements
0	1,148
1	3,351
2	752
3	563
4	1,405
5	852
6	2,812
7	1,128
8	937
9	526

As a result, we have generated a backward model, which combines: a) the (data-driven) 16-dimensional DNN representation vector, belonging to one of 20 cluster centres (10 from each category) and b) the 8-dimensional (semantic) respective input vector of explanatory variables. Whenever a new case appears, the explanatory variables are collected and provided as input to the forward model, implemented through a trained DNN. The latter predicts whether the subject needs to be conveyed.

Moreover, the backward model classifies the respective latent variables extracted from the DNN to the nearest cluster center, thus verifying the DNN-based prediction in an efficient and transparent way. We have validated the ability of the generated backward model - consisting of the 20 cluster centers - to provide the same prediction accuracy as the DNN over the data test set. At the same time, the backward model presents the explanatory variables of the closest cluster center and all related semantic information.

4. Discussion

We have designed convolutional neural networks that were able to outperform other machine learning methods (by approximately 5%), generating a forward model that achieved an accuracy of 80%, ranging from 67% for the conveyance class to 82% for the non-conveyance one. The difference in class prediction accuracy is due to the imbalance in the number of cases between the two classes (about 5 times higher in non-conveyance cases).

In our future work we intend to increase the size of the data set especially with reference to conveyance cases. We have been also applying Generative Adversarial Network (GAN) models [14] for data augmentation. We plan to validate the data obtained through GANs in collaboration with our medical experts, so as to extend the derived forward and backward models with such data as well.

The derived backward model has been able to provide the same prediction accuracy as the DNN model from which it emerged in a computationally efficient, one shot classification approach. In addition, it informs the decision makers of the way that similar approaches

have been treated in the past - by linking the current case to one the cases corresponding to the representative cluster centers, combining data-driven representations and semantic information related to the explanatory variables.

By examining the values of explanatory variables corresponding to the 20 cluster centres in Tables 3 and 4, we can deduce various conclusions. In the conveyance category, the majority of cases refer to subjects with low consciousness (a value of 0) and/or a high value of destination condition (equal or higher to 0.2). Two clusters, with such characteristics, include more than 60 % of the conveyance cases. Some other cases refer to problems in vital signal measurements, such as the BP systolic and BP diastolic values. In one case, a cluster with 7 cases, refers to lack of vital signal values. On the other hand, in the non-conveyance category, we see clusters with high values of consciousness (0.5, or 1.0 values), and/or low values (such as 0) of the destination condition.

Apart from the above cases, where semantic justification can be inferred, there are many other cases, with non-significant differences among the respective values; it is in such cases, where data-driven knowledge, e.g., how former cases have been treated, can be supportive for future decision making.

In our future work we plan to formulate the extracted semantic information through a knowledge representation methodology [24], interweaving it with the machine learning framework [25,26] and attention models [27], extending former work of ours.

We also plan to increase the number of explanatory variables, including more clinical recordings, as well as subject profiling and geographical information. Currently there are missing values, which have not been yet handled, in the data set.

Moreover, we will extend the number of response variables. The need of ambulance can be split in "treated and transported", "treated and discharged", "patient refused transport", "deceased and transported"; the no need case can be split in "no treatment required", "referred to primary care", "referred to other", "own transport".

Our findings will inform development of future interventions and/or strategies to reduce unnecessary hospital admissions following emergencies in care homes, for example through better responses, or pathways following acute episodes of care.

Author Contributions: Conceptualization, Niro Siriwardena and Stefanos Kollias; Data curation, Niro Siriwardena; Formal analysis, Dimitrios Kollias; Investigation, Miao Yu, Dimitrios Kollias and Stefanos Kollias; Methodology, Miao Yu and Dimitrios Kollias; Software, James Wingate; Validation, Niro Siriwardena; Writing – original draft, Dimitrios Kollias and Stefanos Kollias.

References

1. Burke, R.E.; Rooks, S.P.; Levy, C.; Schwartz, R.; Ginde, A.A. Identifying potentially preventable emergency department visits by nursing home residents in the United States. *Journal of the American Medical Directors Association* **2015**, *16*, 395–399.

2. Briggs, R.; Coughlan, T.; Collins, R.; O'Neill, D.; Kennelly, S. Nursing home residents attending the emergency department: clinical characteristics and outcomes. *QJM: An International Journal of Medicine* **2013**, *106*, 803–808.

3. Hsiao, C.J.; Hing, E. Emergency department visits and resulting hospitalizations by elderly nursing home residents, 2001–2008. *Research on Aging* **2014**, *36*, 207–227.

4. Ashcraft, A.S.; Champion, J.D. Nursing home resident symptomatology triggering transfer: avoiding unnecessary hospitalizations. *Nursing research and practice* **2012**, *2012*.

5. Abrahamson, K.; Bernard, B.; Magnabosco, L.; Nazir, A.; Unroe, K.T. The experiences of family members in the nursing home to hospital transfer decision. *BMC geriatrics* **2016**, *16*, 1–7.

6. Stephens, C.; Halifax, E.; Bui, N.; Lee, S.J.; Harrington, C.; Shim, J.; Ritchie, C. Provider perspectives on the influence of family on nursing home resident transfers to the emergency department: crises at the end of life. *Current gerontology and geriatrics research* **2015**, *2015*.

7. Laging, B.; Ford, R.; Bauer, M.; Nay, R. A meta-synthesis of factors influencing nursing home staff decisions to transfer residents to hospital. *Journal of advanced nursing* **2015**, *71*, 2224–2236.

8. Lamb, G.; Tappen, R.; Diaz, S.; Herndon, L.; Ouslander, J.G. Avoidability of hospital transfers of nursing home residents: perspectives of frontline staff. *Journal of the American Geriatrics Society* **2011**, *59*, 1665–1672.

9.   McGregor, M.J.; Abu-Laban, R.B.; Ronald, L.A.; McGrail, K.M.; Andrusiek, D.; Baumbusch, J.; Cox, M.B.; Salomons, K.; Schulzer, M.; Kuramoto, L. Nursing home characteristics associated with resident transfers to emergency departments. *Canadian Journal on Aging/La Revue canadienne du vieillissement* **2014**, *33*, 38–48.

10. Thomas, K.S.; Dosa, D.M. Interviews with Australian nursing home managers reveal that various factors are associated with the decision to transfer a resident to hospital, including two factors not identified in previous research: advance care planning and support from local health services. *Evidence-based nursing* **2012**, *15*, 123–124.

11. Fan, C.; Keating, T.; Brazil, E.; Power, D.; Duggan, J. Impact of season, weekends and bank holidays on emergency department transfers of nursing home residents. *Irish Journal of Medical Science (1971-)* **2016**, *185*, 655–661.

12. Wang, H.E.; Shah, M.N.; Allman, R.M.; Kilgore, M. Emergency Department Visits by Nursing Home Residents in the U nited S tates. *Journal of the American Geriatrics Society* **2011**, *59*, 1864–1872.

13. Letzner, J. Analysis of Emergency Medical Transport Datasets using Machine Learning, 2017.

14. Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. *Deep learning*; Vol. 1, MIT press Cambridge, 2016.

15. Kollias, D.; Tagaris, A.; Stafylopatis, A.; Kollias, S.; Tagaris, G. Deep neural architectures for prediction in healthcare. *Complex & Intelligent Systems* **2018**, *4*, 119–131.

16. Wingate, J.; Kollia, I.; Bidaut, L.; Kollias, S. Unified deep learning approach for prediction of Parkinson's disease. *IET Image Processing* **2020**, *14*, 1980–1989.

17. Kollias, D.; Yu, M.; Tagaris, A.; Leontidis, G.; Stafylopatis, A.; Kollias, S. Adaptation and contextualization of deep neural network models. 2017 IEEE symposium series on computational intelligence (SSCI). IEEE, 2017, pp. 1–8.

18. Min, E.; Guo, X.; Liu, Q.; Zhang, G.; Cui, J.; Long, J. A survey of clustering with deep learning: From the perspective of network architecture. *IEEE Access* **2018**, *6*, 39501–39514.

19. Yang, B.; Fu, X.; Sidiropoulos, N.D.; Hong, M. Towards k-means-friendly spaces: Simultaneous deep learning and clustering. international conference on machine learning. PMLR, 2017, pp. 3861–3870.

20. Arthur, D.; Vassilvitskii, S. k-means++: The advantages of careful seeding. Technical report, Stanford, 2006.

21. Cortes, C.; Vapnik, V. Support-vector networks. *Machine learning* **1995**, *20*, 273–297.

22. Ho, T.K. The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence* **1998**, *20*, 832–844.

23. Haykin, S.; Network, N. A comprehensive foundation. *Neural networks* **2004**, *2*, 41.

24. Kollia, I.; Glimm, B.; Horrocks, I. Query answering over SROIQ knowledge bases with SPARQL. Proceedings of the International Workshop on Description Logic. Citeseer, 2011, Vol. 11.

25. Wallace, M.; Maglogiannis, I.; Karpouzis, K.; Kormentzas, G.; Kollias, S. Intelligent one-stop-shop travel recommendations using an adaptive neural network and clustering of history. *Information Technology & Tourism* **2003**, *6*, 181–193.

26. Kollias, D.; Marandianos, G.; Raouzaiou, A.; Stafylopatis, A.G. Interweaving deep learning and semantic techniques for emotion analysis in human-machine interaction. 2015 10th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP). IEEE, 2015, pp. 1–6.

27. Rapantzikos, K.; Tsapatsoulis, N.; Avrithis, Y.; Kollias, S. Bottom-up spatiotemporal visual attention model for video analysis. *IET Image Processing* **2007**, *1*, 237–248.