Clustering-Aided Approach for Predicting Patient Outcomes with Application to Elderly Healthcare in Ireland

Mahmoud Elbattah, Owen Molloy
National University of Ireland Galway
m.elbattah1@nuigalway.ie, owen.molloy@nuigalway.ie

Abstract
Predictive analytics have proved promising capabilities and opportunities to many aspects of healthcare practice. Data-driven insights can provide an important part of the solution for curbing rising costs and improving care quality. The paper implements machine learning techniques in an attempt to support decision making in relation to elderly healthcare in Ireland, with a particular focus on hip fracture care. We adopt a combination of unsupervised and supervised learning for predicting patient outcomes. Initially, elderly patients are grouped based on the similarity of age, length of stay (LOS) and elapsed time to surgery. Using the K-Means algorithm, our clustering experiments suggest the presence of three coherent clusters of patients. Subsequently, the discovered clusters are utilised to train prediction models that address a particular cluster of patients individually. In particular, two machine learning models are trained for every cluster of patients in order to predict the inpatient LOS, and discharge destination. The developed models are claimed to make predictions with relatively high accuracy. Furthermore, the potential usefulness of the clustering-guided approach of prediction is discussed in general.

Keywords
Machine Learning; Unsupervised Learning; Supervised Learning; Clustering; K-Means; Elderly Healthcare.

1. Introduction

Elderly-driven care continues to become a central issue for healthcare systems owing to the growing trend of population ageing worldwide, and especially in developed countries. Likewise in Ireland, population ageing has been increasing in a rapid pace. As reported by Ireland’s Health Service Executive (HSE), the increase in the number of people aged over 65 is approaching 20,000 per year (HSE 2014). As a result, there is an imperative need for an insightful planning in order to become abreast of the prospective demographic shift. In this respect, machine learning presents as a practical method for building prediction models that can support the planning of healthcare services.

The focus of the study is centralised around the care scheme of hip fracture for elderly patients in Ireland. In particular, the study aimed to build machine learning models that can make predictions on patient outcomes in terms of inpatient length of stay (LOS) and discharge destination.

Patient outcomes were described as the end result of care, or a measurable change in the health status or behaviour of patients (Harris 1991). According to (Nolan and Mock 2000), there are four categories of patient outcomes including: i) Clinical outcomes, ii) Functional outcomes, iii) Financial outcomes, and iv) Perceptual outcomes. The clinical outcomes mainly represent the patient’s response to medical interventions. The functional outcomes measure the improvement of a patient’s physical functioning. Financial outcomes are used to assess the efficient usage of resources. The perceptual outcomes might be the most intangible set of outcomes, which can represent a patient's satisfaction with care received and its providers. The scope of patient outcomes addressed by the paper can be considered to fall within the clinical and financial categories.

The study initially started with the aim of grouping patients from a data-driven perspective. In this regard, unsupervised clustering using the K-Means algorithm was applied. The discovered clusters were then utilised to train the prediction models as delineated at the following sections in detail.

2. Background: Elderly Care and Hip Fractures in Ireland

This section aims to deliver a concise review of the application domain endorsed by the study. As alluded to earlier, the study focused on the care scheme of hip fracture in Ireland. In particular, we were concerned with elderly patients (aged 60 and over) who undergo hip fracture care.

Hip fractures were considered for a set of plausible reasons. First, hip fracture care can represent an adequate exemplar of elderly healthcare schemes. In this regard, plentiful studies (e.g. (Cooper, Campion, and Melton 1992), (Melton, 1996), and (Gullberg, Johnell, and Kanis...
identified hip fractures as exponentially increasing with age, despite the variation of rates from country to another. In Ireland, around 3,000 people sustain hip fractures annually (Ellanti et al. 2014), which will unavoidably increase owing to population ageing. Second, the care delivery of hip fractures is of a considerable importance in Ireland, whereas the HSE identified hip fractures as one of the most serious injuries resulting in lengthy hospital admissions and high costs (HSE 2008). Typically, the cost of treating a hip fracture was estimated around €12,600 (HSE 2008). Equally important, more than two-thirds of patients can spend prolonged periods in long-stay care facilities after hospital discharge (Ellanti et al. 2014).

For all these reasons, hip fractures can represent a major concern to healthcare in Ireland. Predictive analytics can therefore serve as a key enabling factor for healthcare executives through providing evidence-based directions for future strategies.

### 3. Significance of the Study

Generally, the inpatient LOS and discharge destination both represent a pivotal importance within healthcare schemes. According to numerous studies (e.g. (O’Keefe, Jurkovich, and Maier 1999), (Englert, Davis, and Koch 2001), and (Guru et al. 2005)), the LOS was suggested as a significant measure of patient outcomes. While from a strategic perspective, the LOS was recognised as a valid proxy to measure the consumption of hospital resources (e.g. (Millard, Peter, and Sally 1994), (Faddy and McClean 1999), (Marshall and McClean 2003), and (Fackrell 2009)). Further, the LOS was reported as the main component of the overall cost of hip fracture care (Johansen et al. 2013.).

Similarly, predicting discharge destination has a strategic importance in order to estimate the needed capacity of long-stay care facilities such as nursing homes for example. (Brasel et al. 2007) also demonstrated that there is an evident relationship between the inpatient LOS and discharge destination.

In light of that, the study is claimed to carry potential benefits as follows. From a practical standpoint, the developed machine learning models are capable of predicting LOS and discharge destination with relatively high accuracy. Moreover, the discovery of coherent clusters of patients can extend opportunities to answer questions, or raise further unposed questions such as:

- How do patient clusters vary with respect to specific characteristics such as age, gender or fragility history for example?
- How do patient clusters vary with respect to patient outcomes in terms of LOS and discharge destination?
- Is there a possible correlation between patient outcomes and other care-related factors, such as “Time to Surgery” for example?

In a broader context, we believe that the study can be considered within other comparable machine learning problems in healthcare. Specifically, the combined unsupervised/supervised approach adopted by the study can arguably draw further attention to the significance of patient clustering while dealing with prediction-related problems. In this respect, the study evidently demonstrated that the weight of features used for training our models remarkably varied from a cluster of patients to another.

### 4. Approach Overview

The study adopted an approach that comprised a combination of unsupervised and supervised learning conducted over two stages. Initially, the study aimed to explore the potential presence of coherent clusters of patients. The K-Means algorithm was used to discover the patient groups from a mere data-driven perspective. This stage realised the discovery of three well-separated patient clusters as explained through Section 3.2 - 3.5.

The second stage included the development of machine learning models that can make predictions on patient outcomes. The prediction models were developed with respect to every cluster of patients separately. Specifically, a regression model was trained for predicting the LOS, and another binary classifier for predicting discharge destination. We availed of the Microsoft Azure Machine Learning (Barga, Fontama, and Tok 2015) to train and test the prediction models. Figure 1 illustrates the adopted approach.

![Figure 1. Approach overview. The approach is described based on the existence of three patient clusters, which is explained in detail at Section 3.6.](image-url)
5. Methodology

5.1 Data Description

The main source of the data used by the study is the Irish Hip Fracture Database (IHFD) (NOCA 2012). The IHFD repository is the national clinical audit developed to capture care standards and outcomes for hip-fracture patients in Ireland. Decisions to grant access to the IHFD data are made by the National Office of Clinical Audit (NOCA).

We acquired a dataset extracted from the IHFD repository. The dataset included records about elderly patients aged 60 and over. The data comprised 4,773 records over two years, from January 2013 to December 2014. Figure 2 plots a histogram of the age distribution within the dataset, while Figure 3 shows the percentages of male and female patients. It is worth mentioning that a particular patient may be related to more than one record in case of recurrent fractures. However, we were unfortunately not able to ascertain the proportion of recurrent cases because patients had no unique identifiers, and records were completely anonymised for privacy purposes.

The dataset records contained ample information about a patient’s journey from admission to discharge. Specifically, a typical patient record included 38 data fields such as gender, age, type of fracture, date of admission and LOS. A thorough explanation of the data fields was available via the official data dictionary (HPO 2012). Initially, we explored the variables that can serve as features for training our machine learning models. Based on our intuition, many irrelevant variables could be obviously excluded (e.g. Admission/Discharge Time). Table 1 lists the variables that were initially considered as possible features for the clustering and prediction models as well.

<table>
<thead>
<tr>
<th>Variables Explored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Hospital</td>
</tr>
<tr>
<td>Discharge Code</td>
</tr>
<tr>
<td>Residence Area</td>
</tr>
<tr>
<td>Admission Source</td>
</tr>
<tr>
<td>Hospital Transferred To</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Admission Trauma Type</td>
</tr>
<tr>
<td>Pre-Fracture Mobility</td>
</tr>
<tr>
<td>Frailty History</td>
</tr>
<tr>
<td>Multi-Rehabilitation Assessment</td>
</tr>
</tbody>
</table>

5.2 Data Preprocessing

This section describes the data preprocessing phase conducted prior to training our machine learning models including the clustering and prediction models. The preprocessing procedures included: i) Removing outliers, ii) Feature scaling, and iii) Extraction of features that are considered as indicators of care quality. The preprocessing was implemented using the R language.

The existence of data anomalies is largely unavoidable. A data anomaly was defined as an observation that appears to be inconsistent with the remainder of the dataset (Hodge, and Austin 2004), or more generally as any data that is unsuitable for the intended use (Sarsfield 2009). In our case, data anomalies were mainly outliers and data imbalances as explained in the following sections.

A. Outliers Removal

According to (HSE-NOCA, 2014), the mean and median LOS for hip-fracture patients were reported as 19 and 12.5 days respectively. Therefore, we considered only the samples whose LOS were no longer than 60 days in order to prevent the odd influence of outliers. The excluded outliers represented approximately 5% of the overall dataset. Figure 4 plots a histogram of the LOS used to identify outliers.

Figure 2. The distribution of patients’ age within the dataset.

Figure 3. The proportions of male and female patients within the dataset.

Figure 4. Histogram and probability density of the LOS variable. The outliers can be obviously observed when the inpatient stay becomes longer than 60 days.
B. Feature Scaling
Feature scaling is a central preprocessing step in machine learning in case that the range of features values varies widely. Several studies (e.g. (Visalakshi, and Thangavel 2009) and (Patel, and Mehta 2011)) argued that large variations within the range of feature values can affect the quality of computed clusters.

In this regard, the features of the raw data were rescaled to constrain dataset values to a standard range. The min-max normalisation method was used, and every feature was linearly rescaled to the [0, 1] interval. The values were transformed using the formula below:

\[
z = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)
\]

C. Feature Extraction
In a report (BOA 2007) published by the British Orthopaedic Association, six standards for hip fracture care were emphasised. Those standards generally reflect good practice at key stages of hip fracture care including:

- All patients with hip fracture should be admitted to an acute orthopaedic ward within 4 hours of presentation.
- All patients with hip fracture who are medically fit should have surgery within 48 hours of admission, and during normal working hours.

The raw data did not include fields that explicitly captured such standards. However, they could be derived based on the date/time values of patient arrival, admission and surgery. Thus, two new features were extracted named as “Time to Admission (TTA)” and “Time to Surgery (TTS)”. Eventually, only the TTS was included, whereas the TTA contained a significant amount of missing values.

5.3 Clustering Approach
According to (Jain 2010), clustering approaches can be broadly divided into two categories as follows: i) Hierarchical clustering, and ii) Partitional clustering. On one hand, the family of hierarchical clustering algorithms attempt to build a hierarchy of clusters representing a nested grouping of objects and similarity level that change the grouping scope. In this way, clusters can be computed either in an agglomerative (bottom-up) fashion, or divisive (top-down) fashion.

On the other hand, partitional clustering algorithms decompose data into a set of disjoint clusters (Sammut, and Webb 2011). Data is divided into K clusters satisfying that: i) Each cluster contains at least one point, and ii) Each point belongs to exactly one cluster. The K-Means algorithm is an example.

In this paper, we embraced the partitional clustering approach using the K-Means algorithm. The K-Means algorithm has been constantly considered as one of the simplest and most widely used clustering algorithms. The reasons behind the popularity of K-Means are robustly recognised in literature. Examples are the ease of implementation, simplicity, efficiency, and empirical success (Jain 2010) and (Rokach and Maimon 2005).

The K-Means clustering uses a simple iterative technique to group points in a dataset into clusters that contain similar characteristics. Initially, a number of clusters (K) is decided. The algorithm iteratively places data points into clusters by minimizing the within-cluster sum of squares as in Equation (2) below. The algorithm converges on a solution when meeting one or more of these conditions:

- The cluster assignments no longer change.
- The specified number of iterations is completed.

\[
J(C_k) = \sum_{x_i \in C_k} \|x_i - \mu_k\|^2 \quad (2) \quad (\text{Jain 2010})
\]

Where \(\mu_k\) is the mean of cluster \(C_k\), and \(J(C_k)\) is the squared error between \(\mu_k\) and the points in cluster \(C_k\).

5.4 Selected Features
The patients were clustered based on: i) Age, ii) LOS, and iii) Time to surgery (TTS). We considered the numeric features only since the K-Means algorithm is originally applicable to numeric features only, such that a distance metric (e.g. Euclidean distance) can be used for measuring the similarity between data points. However, it is worth mentioning that there are some K-Means extensions that attempted to incorporate categorical features, such as the K-Modes algorithm (Huang 1998).

5.5. Discovering Patient Clusters
The question of “How many clusters are present in the data?” has been a central issue while approaching a clustering task. The clustering experiments were therefore examined using a number of clusters ranging from 2 to 7. In this manner, the appropriate number was decided based on the experimental results. Table 2 summarises the parameters used during the clustering experiments.

With the aid of Principal Component Analysis (PCA), the computed clusters were projected into two dimensions as shown in Figure 5. The sub-figures in Figure 5 represent the output of a clustering experiment using a different number of clusters (K). Initially with K=2, a promising tendency of clustering was indicated, where the data space is obviously separated into two big clusters. Similarly for K=3, the clusters are still well-separated. However, the quality of the clusters started to decline when k=4 upwards. Furthermore, Figure 6 attempts to assess the appropriate number of clusters by plotting the sum of squared distances as demonstrated in Equation 2. In light of that, it turned out that there were three potential clusters of patients that best separated the dataset.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Clusters (K)</td>
<td>2–7</td>
</tr>
<tr>
<td>Centroid Initialisation</td>
<td>Random</td>
</tr>
<tr>
<td>Similarity Metric</td>
<td>Euclidian Distance</td>
</tr>
<tr>
<td>Number of Iterations</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 2. Parameters of the K-Means algorithm.
Figure 5. Clustering experiments with number of clusters (K) ranging from 2 to 7. The clusters are projected based on the two principal components. An inconsistency of colours indicates less separation of clusters. The visualisations were produced using the R package ggplot (Wickham 2009).

Figure 6. Plotting the sum of squared distances within clusters.

5.6 Preprocessing: Tackling Data Imbalances
This section represents the beginning of the second stage of our methodology, where a supervised learning is utilised to train models in order to predict the intended patient outcomes. The training dataset has now been divided into three partitions representing the discovered clusters. The clusters inherently suffered from imbalanced distributions, where a particular set of values predominated. The imbalance issue was pronounced for both of the LOS and discharge destination labels. Figure 7 provides an example of the imbalance within the LOS values.

Figure 7. The LOS histogram. The sub-figures represent LOS distributions within the three discovered clusters of patients. The imbalance of data samples can be observed in the three clusters.

(Galar et al. 2012) identified two different strategies in order to deal with the imbalance problem including: i) Algorithm-level approach and, ii) Data-level approach such as under-sampling or over-sampling. The over-sampling technique (Japkowicz 2000) was adopted in this paper. The underrepresented samples were resampled at random until they contained relatively sufficient examples.

5.7 Learning Algorithm: Random Forests
The study adopted a unified approach using Random Forests for regression and classification as well. Introduced by Leo Breiman (Breiman 2011), Random Forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. During training phase, all trees are trained independently. While testing, predictions are made through weighted voting on the most confident predicted class. The trees that have a higher prediction confidence will have a greater weight in the final decision of the ensemble. The aggregation of voting can be done by a simple averaging operation as in Equation (3). Random forests were proved to provide high accuracy, robustness to noise and stability (Breiman 2011) and (Liaw and Wiener 2002). In addition, the efficiency of Random Forests was evidently demonstrated in a diversity of applications (e.g. bioinformatics (Jiang et al. 2004), ecology (Prasad, Iverson, and Liaw 2006)).

\[ p(c|v) = \frac{1}{T} \sum_{t=1}^{T} p_t(c|v) \] (3)

Where \( p(c|v) \) denotes the posterior distribution obtained by the t-th tree.

Using Random Forests, two models were developed for every cluster of patients. Specifically, the prediction
models included a regression model for predicting the LOS, and a binary classifier for predicting discharge destination. The predicted discharge destination represented either home or a long-stay care facility such as nursing homes. Table 3 presents the parameters used for training the random forests.

| Parameters of random forests models. |
|-----------------|-------|
| Number of Decision Trees | 8     |
| Maximum Depth of Decision Trees | 32    |
| Number of Random Splits per Node | 128   |

5.8 Feature Selection
As mentioned in Section 5.1, the dataset initially contained 38 features, however not all of them were relevant. Intuitively irrelevant feature were simply excluded. In addition, the most important features were decided based on the technique of permutation feature importance (Altmann, Toloşi, Sander and Lengauer 2010). Table 4 presents the set of features used by both models.

<table>
<thead>
<tr>
<th>Prediction Model</th>
<th>Selected Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOS Regression Model</td>
<td>Age, Patient Gender, Fracture Type, Hospital Admitted To, ICD-10 Diagnosis, Fragility History, Time to Surgery</td>
</tr>
<tr>
<td>Discharge Destination Classifier</td>
<td>Age, Patient Gender, Fracture Type, Hospital Admitted To, ICD-10 Diagnosis, Fragility History, Time to Surgery, LOS</td>
</tr>
</tbody>
</table>

It was observed that the importance of features remarkably varied from a cluster to another. For instance, the patient age scored the highest importance within the first cluster while training the LOS regression model. On the other hand, the “Time to Surgery” feature had the highest importance in the third cluster. This can translate into a relative disparity between patient clusters, which obviously reflected on the importance of features used for model training. Table 5 lists the importance of features scored while training the LOS prediction model for every cluster.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Feature Importance Score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Cluster1 0.84, Cluster2 0.49, Cluster3 0.61</td>
</tr>
<tr>
<td>Patient Gender</td>
<td>Cluster1 0.14, Cluster2 0.23, Cluster3 0.18</td>
</tr>
<tr>
<td>Fracture Type</td>
<td>Cluster1 0.38, Cluster2 0.44, Cluster3 0.21</td>
</tr>
<tr>
<td>Hospital Admitted To</td>
<td>Cluster1 0.78, Cluster2 0.93, Cluster3 0.56</td>
</tr>
<tr>
<td>ICD-10 Diagnosis</td>
<td>Cluster1 0.48, Cluster2 0.52, Cluster3 0.29</td>
</tr>
<tr>
<td>Fragility History</td>
<td>Cluster1 0.44, Cluster2 0.10, Cluster3 0.09</td>
</tr>
<tr>
<td>Time To Surgery</td>
<td>Cluster1 0.15, Cluster2 0.27, Cluster3 0.64</td>
</tr>
</tbody>
</table>

6. Experimental Results
The predictive models were tested using a subset from the dataset described in Section 3.1. The randomly sampled test data represented approximately 40% of the overall dataset. The prediction error of each model was estimated by applying 10-fold cross-validation. Table 6 provides evaluation metrics for the LOS regression models, while Figure 8 shows the ROC curves of the discharge destination classifiers. Furthermore, Table 7 presents the measures of precision, recall and accuracy of the classifiers.

<table>
<thead>
<tr>
<th>LOS Predictors</th>
<th>Relative Absolute Error (%)</th>
<th>Relative Squared Error (%)</th>
<th>Coefficient of Determination (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>0.24</td>
<td>0.14</td>
<td>0.86</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>0.30</td>
<td>0.19</td>
<td>0.81</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>0.28</td>
<td>0.18</td>
<td>0.82</td>
</tr>
</tbody>
</table>

(a) AUC=0.851 (b) AUC=0.739 (c) AUC=0.863

Figure 8. The average 10-fold cross-validation accuracy of the discharge destination classifiers. Each sub-figure represents the accuracy of a binary classifier trained for a specific cluster of patients.

<table>
<thead>
<tr>
<th>Discharge Destination Classifiers</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>0.821</td>
<td>0.876</td>
<td>0.810</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>0.733</td>
<td>0.767</td>
<td>0.739</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>0.809</td>
<td>0.857</td>
<td>0.811</td>
</tr>
</tbody>
</table>

7. Comparative Analysis
It was important to evaluate the clustering-aided approach compared to other simpler approaches. In this respect, this section implements a comparative analysis to assess the magnitude of improvement in prediction accuracy using our approach against non-clustering-based predictions.

To achieve this, we developed two additional models including a regression model for predicting the LOS, and a binary classifier for predicting discharge destination. Both models were trained using the same learning algorithm (i.e. Random Forests), and parameters as mentioned earlier in Section 5.7. However, the new
models were trained and tested with respect to the whole dataset without considering patients clusters. The comparative analysis clearly showed that the clustering-aided prediction is more accurate in both cases of predicting LOS and discharge destination. Figure 9 and Figure 10 demonstrate the comparative analysis results.

![Figure 9](image1.png)

**Figure 9. LOS Prediction Accuracy: Clustering–aided predictions against non-clustering. The clustering-aided accuracy is computed as the median of the three models trained for patient clusters.**

![Figure 10](image2.png)

**Figure 10. Discharge Destination Prediction Accuracy: Clustering–aided predictions against non-clustering. The clustering-aided accuracy is computed as the median of the three models trained for patient clusters.**

8. Related Work

The application of machine learning techniques within the context of healthcare was largely applied. The literature contains plentiful studies that utilised machine learning-based methods to predict the LOS, or similar patient outcomes (e.g. (Liu et al. 2006), (Hachesu, Ahmadi, Alizadeh, and Sadoughi 2013), (Azari, Janeja, and Mohseni 2012), and (Morton et al. 2014)). For instance, decision trees and Naive Bayesian classifiers were trained to a geriatric hospital dataset in order to predict the LOS (Liu et al. 2006). Likewise, (Morton et al. 2014) conducted a comparison between supervised machine learning techniques for predicting short-term LOS for diabetic patients.

In pertinent respects, a number of recent studies addressed hip fracture care. For instance, some studies (e.g. (Marufu et al. 2016), (Nijmeijer et al. 2016), and (Karres, Heesakkers, Ultee, and Vrouwenraets 2015)) focused on predicting patient outcomes in terms of mortality. (Marufu et al. 2016) developed a logistic regression model aiming to predict the 30-day mortality after hip fracture surgery. The study utilised a dataset obtained from the UK’s National Hip Fracture Database (Currie et al. 2012).

9. Study Limitations

We acknowledge the limitations of the study as follows:

- The patients were clustered based on a mere data-driven standpoint. It should be taken into account that adding a clinical perspective (e.g. diagnosis, procedures) may group patients differently.
- The IHFD dataset included patients admitted to public hospitals only.

10. Conclusions

The combined unsupervised/supervised approach adopted by the study can draw more attention to the significance of patient clustering while dealing with prediction-related problems. Specifically, the clustering-aided approach yielded further concerns that can contribute to improving the accuracy of predicting patient outcomes.

First, the study evidently demonstrated a pronounced variance of the importance of selected features from a cluster of patients to another. Accordingly, the clustering-aided approach can arguably make finer-grained predictions for patients who naturally share a common set of characteristics, or experienced similar outcomes. In this regard, our approach also proved to provide more accurate predictions compared to non-clustering-based predictions. In a broader context, we believe that the study can be considered within other comparable machine learning problems in healthcare.
References


