

Predicting 30-day Readmission in Heart Failure using Machine Learning Techniques

Jon Kerexeta^{1,2}, Arkaitz Artetxe^{1,2}, Vanessa Escolar³, Ainara Lozano³ and Nekane Larburu^{1,2}

¹Vicomtech-IK4 Research Centre, Mikeletegi Pasalekua 57, 20009, San Sebastian, Spain

²Biodonostia Health Research Institute, P. Doctor Begiristain s/n, 20014 San Sebastian, Spain

³Hospital Universitario de Basurto (Osakidetza Health Care System), Avda Montevideo 18, 48013, Bilbao, Spain

Keywords: Heart Failure, Machine Learning, Hospital Readmission, Risk Prediction, Classification.

Abstract: Heart Failure (HF) is a syndrome that reduces patients' quality of life, and has severe impacts on healthcare systems worldwide, such as the high rate of readmissions. In order to reduce the readmissions and improve patients' quality of life, several studies are trying to assess the risk of a patient to be readmitted, so that taking right actions clinicians can prevent patient deterioration and readmission. Predictive models have the ability to identify patients at high risk. Henceforth, this paper studies predictive models to determine the risk of a HF patient to be readmitted in the next 30 days after discharge. We present two different approaches. In the first one, we combine unsupervised and supervised classification and achieved AUC score of 0.64. In the second one, we combine decision tree and Naive Bayes classifiers and achieved AUC score of 0.61. Additionally, we discover that the results improve when training the predictive models with different readmission's threshold outcome, reaching the AUC score of 0.73 when applying the first approach.

1 INTRODUCTION

Heart Failure (HF) is a clinical syndrome caused by a structural and/or functional cardiac abnormality. It results in a reduced cardiac output (i.e. inability of the heart to pump the blood in the required amounts to satisfy the requirements of the metabolism), and/or elevated intracardiac pressures at rest or during stress. Moreover, HF is associated with a decreased quality of life which reduces physical and mental activity (Ponikowski, 2016).

The prevalence of HF depends on the definition applied, but it is approximately 1–2% of the adults in developed countries, rising to $\geq 10\%$ among people >70 years of age (Ponikowski, 2016). Hence, due to the population's aging, it is expected an increasing of HF patient's number in the future. Furthermore, HF patients often readmit after the discharge, with 56.6% of annual readmission (Maggioni, 2016), which result on high expenses for healthcare systems. For instance, Maggioni et al. (Maggioni, 2016) estimated that €11,867 were spent annually per HF patient, in Italy during 2008-2012.

Considering the expected increment of HF patient's number and the cost associated to each patient, HF will potentially become a big issue in

coming decades unless some actions are taken. In this context, there is a growing interest in reducing the readmission rates.

In this paper we focus on the risk assessment of HF patients' readmission using machine learning techniques, so that this information could help clinicians on managing their patients best by giving a closer follow-up to those patients with higher risk. This way, readmission rate can be potentially reduced, improving the quality of life of HF patients.

This paper is structured as follows: Section 2 – *State of the Art* summarizes how the HF readmission risk is assessed in the literature. Section 3 – *Dataset* presents the dataset applied and Section 4 – *Proposed Methods*, describes the classifiers proposed in this study. Section 5 – *Results*, provides the results of each proposed method. Finally, in Section 6 – *Conclusion and Future Work*, we discuss the conclusions and future work that will follow this study.

2 STATE OF THE ART

There are a plethora of studies on readmission risk prediction modelling. This section presents a brief

summary of studies in the context of HF patient readmission prediction (see Table 1).

Mortazavi et al. (Mortazavi, 2016) compare prediction techniques in HF readmission: logistic regression (LR), Poisson regression (PR), random forest (RF) and, boosting and random forest combined hierarchically with support vector machines (SVM). Their dataset have 977 patients and 236 attributes. They achieve 0.543 AUC value using LR and 0.615 using boosting. They improve the results using a readmission threshold of 180 days: 0.669 AUC using RF and 0.678 using boosting.

Zolfaghar et al. (Zolfaghar, 2013) investigate the readmission risk in HF, defining the outcome of 30-days readmission. They focus in the fact that data is not well balanced (i.e., many more no readmissions than readmissions, with a proportion of 1:5.7). To solve this problem, they make more than one classifier in different layers, achieving sensitivity (Se) of 0.31 and specificity (Sp) of 0.81.

Zheng et al. (Zheng, 2015) create several classifiers to estimate if a patient with HF would readmit within 30 days. They use neural networks (NN), SVM with different kernels and RF. They have a dataset of 1641 patients that had an admission because of HF, and of those, 316 patients readmitted within 30-days because HF. The best result that they report were obtained with particle swarm optimization-SVM, achieving a Se = 0.08, Sp = 0.97 and accuracy of 0.78.

Meadem et al. (Meadem, 2013) present a study of readmission prediction within 30-days (all cause), in patients with HF. They focus on the feature extraction of the data: attribute selection (with chi-square and stepwise), missing value imputation (with clustering) and data balancing (with over-sampling and under-sampling). They compare the performance of three different classifiers LR, SVM and NB. Their dataset is composed of 8,600 patients and 49 attributes, after an attribute reduction/selection process. The best result reported is AUC= 0.64, with stepwise attribute selection, clustering missing value imputation, over-sampling data balancing and SVM classifier.

Krumholz et al. (Krumholz, 2000) try to predict whether after HF admission, patients are going to readmit within 6 month. They use random survival forest and Cox regression. They use a sample of 2,176 patients which had a HF admission (if more than one, only the first readmission was considered). They created different risk levels, obtaining the precision value of 0.31 in the best case.

Table 1: Summary of related studies and methodologies.

Author	Data size	Method	AUC
Mortazavi et al.	977	LR, PR and RF	0.687
Zolfaghar et al.	9,770	SVM & NB	0.31*
Zheng et al.	1,641	NN, RF,SVM	0.35*
Meadem et al.	8,600	SVM, NB & LF	0.64
Krumholz et al.	2,176	Cox regression	0.31*
Amarasingham et al.	1,372	Multivariate analysis	0.45*
Sudhakar et al.	1,046	RR score	0.61
Artetxe et al.	119	RF & SVM	0.65

*Precision or Sensitivity values

Amarasingham et al. (Amarasingham, 2010) estimated the 30-days readmission because of HF and mortality risk using multivariate analysis. They applied a dataset comprised of 1372 HF patients, of which 331 readmitted and 43 died within 30-days after discharge. After the multivariate analysis of the data, they got a precision of 0.456 in the best quintile that they created.

Sudhakar et al. (Sudhakar, 2015) develop an all-cause readmission risk score (RR score) in people with HF. They employ a sample of 1,046 admissions (from 712 patients) because of HF, and of those, 369 readmitted within 30-days (all cause). They do a multivariate analysis to get the RR score and 0.61 is the best AUC value that they achieve.

Artetxe et al. (Artetxe, 2017) built classifiers to identify HF patients that have high risk of readmission caused by HF. They focus on the feature extraction. They use filter, wrapping and embedding methods to extract the features, and then, RF and SVM classifiers. The best performance they achieve, using a dataset of 119 cases, is AUC=0.647 with Wrapping extraction method and linear SVM classifier.

3 DATASET

The public hospital OSI Bilbao-Basurto (Osakidetza), located in Basque Country (Spain), has been gathering HF patients' information from 2014 until 2017. For the present study, the dataset contained a cohort of 231 HF patients. Clinicians have collected baseline data (information collected by a clinician when the patient was diagnosed with HF, Table 2), ambulatory patient monitored data (i.e. information that patients at home collect from three to seven times per week, e.g. heart rate) and patients' admissions information (i.e. information related to admission, e.g. length of stay).

Table 2: Summary of the dataset.

Attribute	Description	Mean±SD
Age	The age of the patient (years)	77.7±11
Sex	The sex of the patient (men/women)	57% men*
Smoker	If the patient smoke, did smoke and now do not, or never has smoked	18% Yes* 59% No* 23% Ex*
LVEF	Left Ventricular Ejection Fraction (%)	42.2±15.3
FirstDiag	Years since first diagnosis	6.15±7.23
Implanted device	If the patient has implanted a device (yes/no)	23% Yes*
Need oxygen	If the patient needs oxygen (yes/no)	6% Yes*
<i>Laboratory</i>		
Urea	Urea (mg/dl)	73.7±37.7
Creatinine	Creatinine (mg/dl)	1.3±0.5
Sodium	Sodium (mEq/L)	140±4.2
Potassium	Potassium (g/dl)	4.27±0.76
Haemoglobin	Haemoglobin (g/dl)	13±10.25
<i>Comorbidities</i>		
Sinus rhythm	If the patient has sinus rhythm (yes/no)	37% Yes*
Comorbidity A.F.	If the patient has atrial fibrillation (yes/no)	57% Yes*
Pacemaker	If the patient has a pacemaker (yes/no)	14% Yes*

*Proportions of the labels

In this study we aim to compare our results with the state of the art by using the baseline information (Table 2) and patients’ admissions. The dataset contains 162 admissions caused by HF decompensations, and of those admissions, 36 were readmissions within 30-days caused by HF.

In addition to the attributes of Table 2, we added the attributes *previous admission* and *seasons* to determine whether these also have an impact on the risk assessment.

The attribute *previous admission* consists in how many admissions the patient had in the previous six months. If the patient did not have any readmission in the previous six months, would be the factor “none”. If the patient had 1 to 2 admissions, would be the factor “some readmissions”. Otherwise, if the patient had more than 2 admissions in the previous six months, would be the factor “many readmissions”. The attribute *seasons* consists in the period at which the patient was discharged:

- *winter*: discharge between 12/01 and 03/01

- *spring*: discharge between 03/01 and 06/01
- *summer*: discharge between 06/01 and 09/03
- *autumn*: discharge between 09/03 and 12/01

This attribute was included since clinicians have noticed by experience that there are more admissions/readmissions in autumn-winter than in summer due to the incidence of respiratory infections, which worsen heart failure. In summer, however, hypotension is more frequent (when extreme temperatures happen).

4 PROPOSED METHODS

In this section we present the two methods that we propose to identify patients at high risk of being readmitted. In *Classifier with Clusters* we combine unsupervised and supervised classifiers. In *Hybrid Tree*, we build a hybrid classifier using decision tree and NB classifiers.

4.1 Classifier with Clusters

This approach combines different machine learning methods to build a classifier. The scheme that it follows is shown in Figure 1, and it is explained in the next sections.

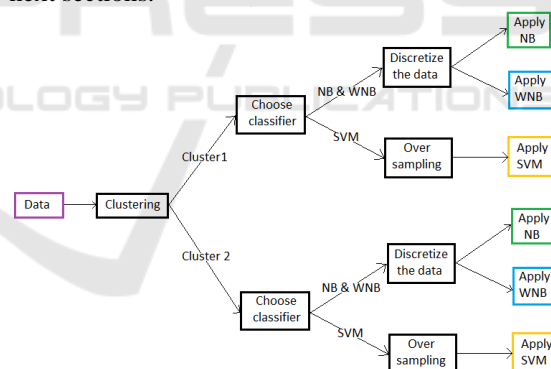


Figure 1: Classifier with Clusters’ scheme.

4.1.1 Clusters

Firstly, we apply *Ward’s* agglomerative hierarchical clustering method (Ward, 1963) with *Manhattan* distance to the dataset, from which we could distinguish two significant clusters (Figure 2). Hence, we work with each cluster separately, building a “specialized” classifier for each one. This way when a new patient risk will be assessed the applied classifier will be the one that works best with similar patients.

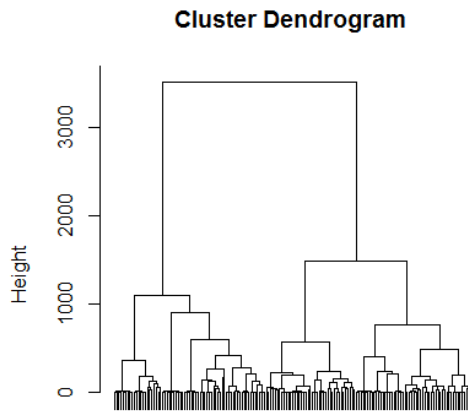


Figure 2: Dendrogram of the data, using *Ward's* method and *Manhattan* distance.

In each cluster we evaluate three classification algorithms, namely SVM (Vapnik, 2013), NB (Murphy, 2006) and weighted Naïve Bayes (WNB) (Zhang, 2004). In order to overcome the class imbalance problem, we do random oversampling, and since NB and WNB need discrete attributes, we apply SVM-based discretization (Section 4.1.2).

4.1.2 Discretization

In order to apply NB and WNB classifiers, the numeric attributes have to be discretized. For that, the optimal cut-point is determined depending on the outcome. To do this, we use a technique based on SVM, similar to one used in Park & Lee (Park, 2009) study, so that results in new binary attributes.

Moreover, we have to find a balance in SVM (“new attributes”) number. With more “new factor attributes” it is possible to have more information, but if the attributes are combined to make SVM classifiers, it is possible to obtain more accurate “new factor attributes”. Hence, we have to decide how to group the attributes when applying SVM classifiers. In this study, *Ward's* agglomerative hierarchical method has been applied to group the attributes.

When two attributes are correlated, if they are grouped the results may be better (separated could be redundant). To determine their relation, we define the distance between the attributes based on the dependence between them:

$$distance(attributes) = 1 - abs(cor(attributes))$$

As a result, the dendrogram shown in Figure 3 is obtained for Cluster 1 using this distance and *Ward's* method. The resulting groups are the following: (1) FirstDiag, and sodium, (2) LVEF and potassium (3) urea, creatinine, haemoglobin and age (Figure 3).

Same procedure is followed in Cluster 2 to group the attributes.

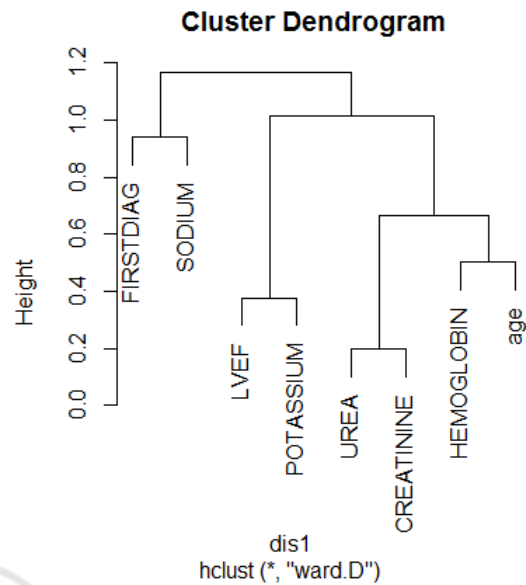


Figure 3: Dendrogram of the attributes in Cluster 1.

Hereafter, SVM based discretization (Park, 2009) is applied in each group.

4.1.3 Weights

Once all the attributes are discretized, it is possible to apply NB classifier to the clusters. But in order to study the impact of WNB, the weights of the attributes have to be determined. For that, in this study the weights are defined as the dependence between attributes and the outcome.

To determine the level of the dependency between attributes and the outcome, *Cramer's V* method (Cramér, 2016) has been applied. This method does not give only information on whether there is a dependence, but also gives the degree of the dependency, which is used as weights for the WNB.

4.2 Hybrid Tree

The clinicians involved in the project consider that classifiers as NB or decision tree are easier to interpret for them than other classifiers since they can be interpreted as rules. Therefore, we develop a classifier that combines both, called *Hybrid Tree*.

This classifier is a hybrid one. Firstly, we use a tree classifier to reduce the data to similar elements. Then, instead of giving the most frequent label, we use a NB classifier to estimate the correct label. This way we take advantage of all the attributes, instead of

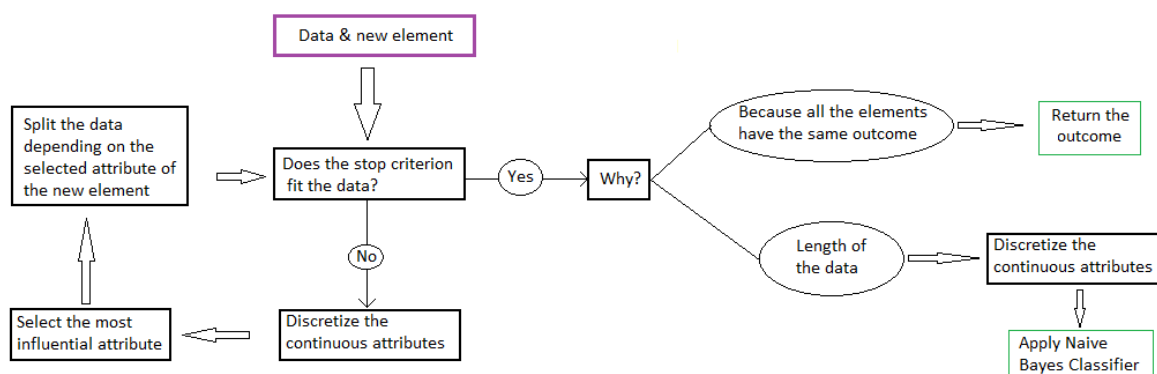


Figure 4: The procedure scheme of the Hybrid Tree.

using only those that the tree requires, and we get more accurate results.

Among all the studies that use similar hybrid tree classifiers, the study of Kohavi et al. (Kohavi, 1996) drew our attention. They compare the performance of hybrid classifier with decision tree and NB classifiers in 29 datasets with different number of elements (100-8,000) and they find that the hybrid one outperforms both (Tree classifier and NB).

4.2.1 Method

We use a recursive method, by which we perform the following steps until a previously defined stopping criterion is met:

We start with the data and the element to classify, and we follow the procedure shown in Figure 4.

- **Stop criterion:** If the data has only one type of label (1) or if the number of cases is too small for further splits (2), we would stop, and (1) return the label or (2) use NB.
- **Discretize:** We discretize all the numeric attributes of the data in each step. We apply SVM to each continuous attribute. Hence, we convert into binary the continuous attributes.
- **The attribute:** We have to choose with which attribute we want to make the split. For that, we calculate the dependency between the discrete attributes and the outcome, and we take the one with the maximum correlation using *Cramer's V* method.
- **Split the data:** We take the cases from the data that have the same selected attribute of the element to classify, and go to *stop criterion*.
- **Naïve Bayes:** If the stop criterion is met because of the length of the data, we apply NB classifier.

4.3 Evaluation Methodology

Due to the limitation on the number of cases from the applied dataset, we use *Leave One Out* (Kearns, 1999) as model validation technique and sensitivity, specificity, precision and area under ROC curve (AUC) metrics (Zou, 2007) to evaluate the models. Owing to the readmission problem, the evaluation metrics are defined as follows:

- **Sensitivity:** among all the readmissions, how many readmissions the classifiers has labelled as readmission
- **Specificity:** among all the no readmissions, how many the classifier has labelled as no readmission
- **Precision:** among all labelled as readmission, how many are correctly identified as such.

5 RESULTS

In this section, we evaluate the proposed classifiers described in Section 4, and additional experiments are made varying the readmission threshold (THR).

5.1 Classifier with Clusters

5.1.1 Clustering

In this subsection, the differences between the clusters built in Section 4.2.1 is studied. Firstly we check whether there is any difference on the attributes' values from each of the clusters, and then we see if the applied clustering method has an influence on the outcome.

The biggest difference between clusters is appreciated with *urea* attribute. In Cluster 2, all the patients have the urea level higher than 80 (the mean

is 113). In Cluster 1 only four patients have the Urea level higher than 80 (the mean is 53). The urea level determines renal function, which is closely related to heart function.

In general, in cluster 2 we detect that most of the attributes have “worse” values associated with worse prognosis – most of them have higher value, except LVEF, where low levels indicate a deterioration. From this, we can suggest that a patient from cluster 2 has higher risk of readmission.

Besides, we have compared the proportions of the discrete attributes of each cluster. We discover that, for example, in Cluster 1 5.3% of the patients need oxygen, while in Cluster 2 27.5% of the patients need oxygen.

There are also significant differences in the outcomes – readmission rate – which is a relevant fact we consider. Figure 5 illustrates Kaplan-Meier curve with the progression of the readmission in each cluster. As presented, Cluster 2 has higher readmission risk compared with those from Cluster 1 (Figure 5).

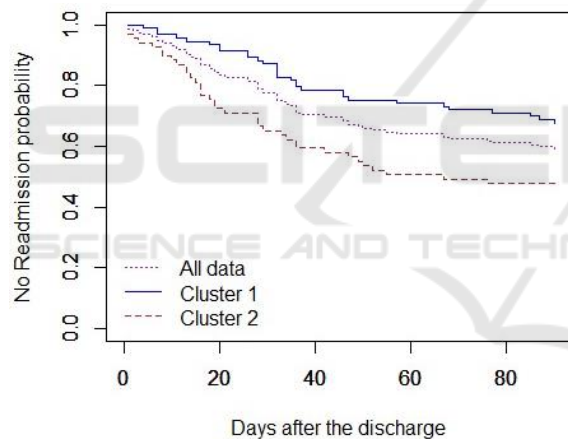


Figure 5: Kaplan-Meier of HF patient’s readmission rate for the whole dataset, Cluster1 and Cluster 2.

5.1.2 Performance of the Classifier

Once the clustering is performed, we apply the classifiers to the clusters and to all the dataset to check whether clustering improves the results.

Table 3: Results we obtained with the classifiers.

	Se	Sp	Precision	AUC
With clusters				
WNB	0.69	0.68	0.39	0.58
NB	0.53	0.75	0.38	0.62
Comb	0.64	0.71	0.40	0.64
SVM	0.67	0.60	0.35	0.62

Without Clusters				
WNB	0.38	0.79	0.34	0.62
NB	0.33	0.88	0.44	0.61
SVM	0.5	0.73	0.35	0.60

As shown in Table 3, the best results are obtained using the combined classifier, i.e. in one cluster WNB and in the other NB. We also get suited results with NB and SVM with clusters. Even if WNB's AUC is lower than others, it has the highest sensitivity. Therefore, we also consider it in our study.

If we compare the classifiers that use clusters with those that do not use clusters, the results are very similar in terms of AUC (Table 3). If we look at Se, classifiers with clusters obtain higher scores, which is very important because of the nature of the problem.

5.2 Hybrid Tree

In Section 4.2, we proposed the Hybrid Tree classifier since clinicians involved consider it easier to interpret the results if the cause is represented as a tree. The results of the *Hybrid Tree* are Se=0.44, Sp=0.82, precision=0.41 and AUC=0.61.

As presented in Section 5.2, similar AUC results are obtained when employing Classifier with Clusters. However, lower Se value is achieved with Hybrid Tree. Therefore, it may be recommended the usage of the first approach from the clinical point of view.

5.3 Readmission-day Threshold

In order to visualize the impact of the readmission day on the obtained results, we visually represent how the implemented classifiers work, using as an example the Kaplan-Meier curves obtained with the Hybrid Tree (Figure 6).

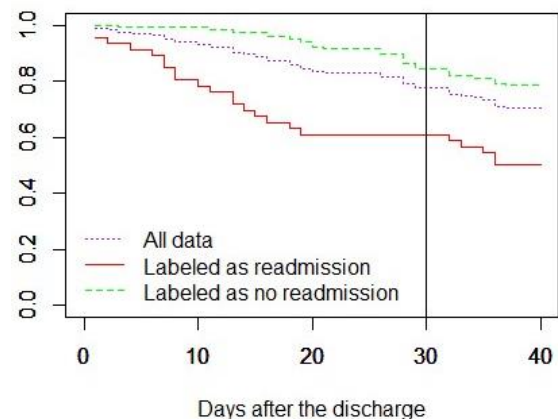


Figure 6: Kaplan-Meier curve of the results of Hybrid Tree.

Figure 6 represents the Kaplan-Meier curve of HF patients dataset (purple) and the curves for those that Hybrid Tree has classified as readmissions (in red), and as no readmission (in green).

The no-readmission curve (in green) is expected to be continuous at zero until the 30th day if the prediction is 100% accurate, but even few of those that are detected as no readmission readmit before 30 days. On the other hand, the readmission curve (in red), should decrease all the way to 0 by the 30th day for perfect prediction. But instead, there is no readmission between approximately day 20-30, and there is a big slope after day 30. Similar behaviour is detected after the 30th THR when plotting Kaplan-Meier curves for Classifiers with Clusters.

This way it is possible to notice from Figure 6 that despite training the Hybrid Tree for 30-days readmission THR, this THR may not be optimal. Notice that this may also depend on the applied dataset.

Furthermore, there is a precedent in the literature (Mortazavi, 2016) that improves remarkably the results training the classifier with 180-days of readmission and testing with 30-days readmission. Therefore, we decided to explore the presented methods when the 30-day readmission THR is modified (Section 5.4.1 and Section 5.4.2).

5.3.1 THR: Classifier with Clusters

Firstly, we check the *Classifier with Clusters* with several THRs using the same THR for training and testing each of them. The best results are obtained with SVM using as THR 35-days (AUC = 0.788).

However, to evaluate the results with the outcome of 30-days readmission (*de facto standard*) we train the classifiers with different THRs and test how they perform with 30-days readmission THR (Table 4).

Table 4: Classifiers with Clusters trained with different readmission days THR and tested with 30-day readmission THR.

Days		30	35	40	60
Se	WNB	0.69	0.72	0.72	0.67
	NB	0.53	0.61	0.61	0.58
	SVM	0.64	0.72	0.72	0.64
Sp	WNB	0.68	0.67	0.67	0.63
	NB	0.75	0.76	0.75	0.72
	SVM	0.64	0.77	0.77	0.66
AUC	WNB	0.58	0.66	0.67	0.65
	NB	0.62	0.64	0.65	0.71
	SVM	0.61	0.73	0.71	0.65

Table 4 presents the improvement of results, where the best results are achieved using SVM

(outcome 35-days), with an AUC = 0.726, and Se = 0.72.

5.3.2 THR: Hybrid Tree

We also test different readmission THRs with Hybrid Tree (Table 5). In this case also the results improve, but the difference is not as high as with Classifier with Clusters.

Table 5: Results of Hybrid Tree, training with different readmission day THR.

Days	30	35	40	60
Se	0.44	0.51	0.54	0.52
Sp	0.82	0.80	0.81	0.72
Precision	0.41	0.51	0.54	0.51
AUC	0.61	0.65	0.65	0.63

6 CONCLUSIONS AND FUTURE WORK

In this study, we present two predictive models to estimate the risk of 30-day readmission in Heart Failure (HF) patients. The first approach combines unsupervised and supervised classifiers (*Classifiers with Clusters*), and the second one, combines decision tree and Naïve Bayes (NB) classifiers (*Hybrid Tree*).

It has been discovered that training the predictive models with different readmission day threshold (higher than 30 days) the results may improve, although it could be related to dataset limitations.

In this context, the best AUC score obtained in this study has been 0.726 with *Classifier with Clusters* (with Support Vector Machines in the clusters), by training it with 35-days readmission THR.

Furthermore, it is also observed that the results substantially improve when the 30-day readmission prediction THR is also extended. For example, when training and testing with 35-days readmission, the result is AUC = 0.788 applying SVM. As discussed, this phenomenon could be due to the dataset size limitations, but relevant to consider.

The AUC values for weighted Naïve Bayes (WNB) and NB are similar. But with WNB the sensitivity values are higher, and with NB, the specificity values are higher. Hence, depending on the problem's nature, we could choose one of the classifiers.

The *Hybrid Tree* classifier performs with an AUC value of 0.65 if the training dataset is considered with 35 or 40 readmission's days (with Se=0.51, Sp=0.80 and Se=0.54, Sp=0.81 respectively). However, the

results of the *Hybrid Tree* may improve over time when the size of the dataset increases (Kohavi, 1996). Due to the nature of the problem the results do not present very high predicting power. Nevertheless, comparing this study with results from the state of the art, the obtained results are satisfactory.

In the future, several actions are planned. First, the presented classifiers will be trained with larger amount of data as new patients are included into the study. In parallel, the ambulatory patient monitored data will be studied to determine whether the presented predictive models could be improved. Next, we aim to build an integrated telemonitoring system that integrates these predictive models to support both clinicians, to manage best the patients, and patients, to empower them in their disease management and prevent potential decompensations. Finally, this system will be tested in a trial study to determine its usability.

REFERENCES

- Ponikowski, P., Voors, A. A., Anker, S. D. et al (2016). 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *European heart journal*, 37(27), 2129-2200.
- Maggioni, A. P., Orso, F., Calabria, S. et al (2016). The real-world evidence of heart failure: findings from 41 413 patients of the ARNO database. *European journal of heart failure*, 18(4), 402-410.
- Mortazavi, B. J., Downing, N. S., Bucholz, E. M. et al (2016). Analysis of machine learning techniques for heart failure readmissions. *Circulation: Cardiovascular Quality and Outcomes*, CIRCOUTCOMES-116.
- Zolfaghar, K., Meadem, N., Teredesai, A., Roy, S. B. et al (2013, October). Big data solutions for predicting risk-of-readmission for congestive heart failure patients. In *Big Data, 2013 IEEE International Conference on* (pp. 64-71). IEEE.
- Zheng, B., Zhang, J., Yoon, S. W. et al (2015). Predictive modeling of hospital readmissions using metaheuristics and data mining. *Expert Systems with Applications*, 42(20), 7110-7120.
- Meadem, N., Verbiest, N., Zolfaghar, K. et al (2013). Exploring preprocessing techniques for prediction of risk of readmission for congestive heart failure patients. In *Data Mining and Healthcare (DMH), at International Conference on Knowledge Discovery and Data Mining (KDD)* (Vol. 150).
- Krumholz, H. M., Chen, Y. T., Wang, Y. et al (2000). Predictors of readmission among elderly survivors of admission with heart failure. *American heart journal*, 139(1), 72-77.
- Amarasingham, R., Moore, B. J., Tabak, Y. P. et al (2010). An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data. *Medical care*, 48(11), 981-988.
- Sudhakar, S., Zhang, W., Kuo, Y. F. et al (2015). Validation of the readmission risk score in heart failure patients at a tertiary hospital. *Journal of cardiac failure*, 21(11), 885-891.
- Artetxe, A., Larburu, N., Murga, N. et al (2017, June). Heart Failure Readmission or Early Death Risk Factor Analysis: A Case Study in a Telemonitoring Program. In *International Conference on Innovation in Medicine and Healthcare* (pp. 244-253). Springer, Cham.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301), 236-244.
- Murphy, K. P. (2006). Naive bayes classifiers. *University of British Columbia*.
- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.
- Park, C. H., & Lee, M. (2009). A SVM-based discretization method with application to associative classification. *Expert Systems with Applications*, 36(3), 4784-4787.
- Cramér, H. (2016). *Mathematical Methods of Statistics (PMS-9)* (Vol. 9). Princeton university press.
- Kohavi, R. (1996, August). Scaling Up the Accuracy of Naive-Bayes Classifiers: A Decision-Tree Hybrid. In *KDD* (Vol. 96, pp. 202-207).
- Kearns, M., & Ron, D. (1999). Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. *Neural computation*, 11(6), 1427-1453.
- Zou, K. H., O'Malley, A. J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 115(5), 654-657.
- Zhang, H., & Sheng, S. (2004, November). Learning weighted naive Bayes with accurate ranking. In *Data Mining, 2004. ICDM'04. Fourth IEEE International Conference on* (pp. 567-570). IEEE.