Autoencoder to Predict Hospital Readmissions for Post-Operated Patients Based on Cardiology Data

Júlia A Ferreira[1,2], Thaynara S Matos[1,2], Pietro F Magaldi[1,2], Jasmine B Nunes[1,2]
Pietro CCO Martins[2], Camila R Moreno[2], Guilherme CM Rebello [2],
Ahmad A Almazloum[1,2], Emely P Silva[1], Anderson R Rocha[1]

[1] Recod.ai Lab, Institute of Computing, University of Campinas, Campinas, Brazil
[2] Heart Institute, Hospital das Clínicas HCFMUSP, University of São Paulo, São Paulo, Brazil

Abstract

Telemonitoring of cardiac patients, particularly in the postoperative period, can be improved through Machine Learning architectures that enable remote clinical supervision. Dimensionality reduction models such as Autoencoders (AEs) are suitable for processing cardiological data in this context. In this study, vital signals collected from smartwatches, such as Heart Rate, Blood Pressure (SBP and DBP), and Peripheral Oxygen Saturation (SpO_2), along with patient history, were used as input for a supervised AE-based binary classification model to predict hospital readmission. Data from 49 postoperative cardiac patients were collected over 30 ± 3 days, with a 9/49 readmission rate. After preprocessing, the data were passed through an AE architecture with dense layers, batch normalization, dropout, and a latent space. A total of 63 input combinations were evaluated across latent space dimensions of 8, 12, 16, 20, and 24. The model classified patients as readmitted or not. Results show that including patient history significantly improves prediction. Cross-validation revealed the best performance for the SBP + History input, with an average F1-score of $83.41\% \pm 3.22$. *These findings highlight the model's potential, although further architectural optimization and larger datasets are needed to ensure robustness and clinical applicability.*

1. Introduction

Data processing of vital signals captured by smartwatches through integrated telemonitoring system can improve the status of cardiac patients after surgery. Recent advances in mobile and wearable technologies offer comfort, ease of use, and show great potential in the detection of cardiac arrhythmias [1]. This type of application represents a promising area for the development of new clinical monitoring tools, where identifying cardiac risks associated with patients can contribute to improving healthcare quality. Therefore, this medical application constitutes an effective contribution to patient care services, opening new possibilities for remote monitoring [2].

Based on data collection via smartwatches, various machine learning techniques can be implemented to detect cardiac anomalies, aiding patient follow-up in a remote monitoring setting. The AutoEncoder (AE) is a model used for dimensionality reduction, that represents a viable technique for anomaly detection, reducing data complexity and capturing irregularities in a lower-dimensional space [3].

Given the potential integration of diverse technologies and data processing techniques, the assimilation of cardiological data and Machine Learning models for anomaly detection and risk identification becomes a promising pathway in the medical field. Therefore, the aim of this study is to develop a supervised AE model using cardiological data captured by smartwatches, along with patient history, to predict hospital readmission. This was performed through dimensionality reduction of the data, proposing a potential support tool for clinical monitoring.

2. Methods

Data from 49 post-operative patients were collected in partnership with InCor – the Heart Institute of the Hospital das Clínicas, University of São Paulo Medical School (FMUSP) – using the Samsung Galaxy Watch 5 [4]. Data collection spanned an average period of 30 ± 3 days, capturing the following vital signals from the smartwatch: heart rate (HR) (both ECG and PPG-based), blood pressure, which is systolic (SBP) and diastolic (DBP), and peripheral oxygen saturation (SpO_2). Additionally, the patient's clinical history was included as an input variable, encompassing demographic data (such as age, ethnicity, and BMI), associated conditions (e.g., diabetes, heart failure, liver disease), as well as vital signal measurements by both the smartwatch and standard medical devices (for HR, SBP, DBP, and SpO_2). Patients were categorized using a binary label indicating emergency room (ER) readmission or not, with 9 readmitted patients and 40 not-readmitted. Other clinical metadata were also recorded, such as the number of days until readmission, total visits, and Manchester Triage Scale (ranging from 1 = most urgent to 5 = least urgent).

The time-series vital signals of interest (HR-ECG, HR-PPG, SBP, DBP, SpO_2) were collected by Web FAPO SI^3 platform [4], which is an internal tool created for this purpose, and were preprocessed using different strategies. Missing values (NaNs) were replaced with zeroes, random values, or the mean of the respective signal data. HR-ECG values of zero were corrected using mean, median, or random values from the same patient. The signals were subsequently under three scenarios: no transformation, Fourier Transform, or Wavelet Transform in previous optimizations tests using Optuna. The best configurations and classifications were related the use of the Wavelet Transform, which was fixed in this study. Finally, all variables were normalized using Min-Max scaling. An example of the signals after preprocessing is shown in Figure 1.

Figure 1. Example for time-series vital signals after the preprocessing

The neural network input vectors were based on these standardized and normalized signal data. The AutoEncoder architecture employed consists of an encoder with multiple layers to compress the input into a latent space, extracting the most informative features at each layer. The decoder mirrors the encoder to reconstruct the input, aiming to minimize reconstruction error. The model uses five dense layers interleaved with batch normalization and dropout layers (30%), a learning rate of 0.1% with the Adam optimizer, and mean squared error (MSE) as the loss function, trained over 200 epochs using TensorFlow. The architecture is shown in Figure 2.

In the patient history, categorical labels were converted into binary values (0 or 1) using the OneHotEncoder module, and only numerical variable were normalized. The latent space was used for classification, containing the most informative features. Multiple latent space dimensions were tested (8, 12, 16, 20, and 24) to evaluate classification performance. To assess the impact of each variable, all 63 possible combinations among the 6 input variables were evaluated. For validation, a 3-fold cross-validation scheme was applied, alternating the train/test groups for each fold. The classification performance was assessed using F1-score, ROC-AUC, recall, and balanced accuracy, with the average with standard deviation across the three folds computed for result analysis.

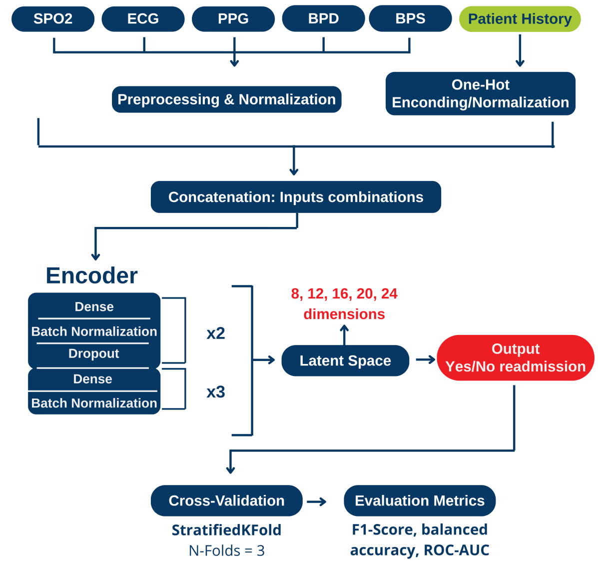

Figure 2. Block diagram of the data processing for time-series vital signals and patient history

3. Results

All 63 combinations of vital signals and patient history were tested across different latent space dimensions. Initially, the metric results corresponding to the average performance across the 3 folds using all concatenated variables (SBP + DBP + HR-ECG + HR-PPG + SpO_2 + History) for different latent space dimensions are presented in Table 1.

It was observed that, for all concatenated variables across varying dimensions, the evaluation metrics yielded similar values. Although the F1-scores obtained were promising overall, they primarily reflected good performance in predicting non-readmitted patients, while all readmitted patients were misclassified. This limitation prompted the evaluation of different input combinations to identify which variables contributed most to patient classification.

Table 1. Average results for all variables (SBP + DBP + HR-ECG + HR-PPG + SpO_2 + history)

Dimension	Mean F1-Score	Balanced Accuracy	ROC-AUC
8	$71.14\% \pm 4.34$	$47.43\% \pm 2.22$	$45.23\% \pm 28.16$
12	$73.35\% \pm 0.88$	$50.00\% \pm 0.00$	$46.88\% \pm 17.16$
16	$72.29\% \pm 2.41$	$48.71\% \pm 2.22$	$47.61\% \pm 5.64$
20	$73.35\% \pm 0.88$	$50.00\% \pm 0.00$	$50.54\% \pm 19.03$
24	$73.35\% \pm 0.88$	$50.00\% \pm 0.00$	$39.74\% \pm 8.97$

The results for the various input combinations are shown in Table 2. Among the top-performing configurations, patient history and blood pressure variables (SBP and DBP) emerged as the most relevant inputs. The inclusion of patient history consistently improved model performance, highlighting its importance in hospital readmission prediction. This improvement is likely due to the presence of clinically relevant information within the patient history, such as records of previous cardiac conditions, which enhance the model's ability to identify scenarios requiring hospital readmission. Heart rate variables (HR-ECG and HR-PPG) also contributed to classification performance, especially when combined with other signals, most notably patient history and blood pressure. Under the current preprocessing pipeline and AutoEncoder (AE) architecture, peripheral oxygen saturation (SpO_2) demonstrated limited impact on model performance.

Furthermore, latent dimensions of 16 and 20 provided the best predictive outcomes. Evaluation metrics demonstrated promising results, particularly balanced accuracy, with averages exceeding 61%. The mean F1-score surpassed 77%, though this metric does not account for class imbalance, specifically, the underrepresented class of ER readmissions, highlighting the need for future adjustments to address imbalance. The ROC-AUC values were also encouraging, averaging above 59%, but showed high standard deviations, indicating the need for further refinement.

The confusion matrices in Figure 3 illustrate the classification performance for the two best combinations across the concatenated test sets.

The confusion matrices indicate satisfactory classification for not-readmitted patients. However, improvements are needed for readmitted cases, as two severe patients (Manchester level 2, very urgent) were misclassified as false negatives. These cases involved atrioventricular block and chest pain/cough. Additionally, other patients classified as Manchester level 3 (urgent) were also incorrectly predicted as not-readmitted. This highlights the need for further model refinement to improve anomaly detection and readmission prediction.

The main limitations and challenges in this binary classification task are primarily associated with the dataset, which contains a limited number of patients, particularly those who returned to the ER, resulting in class imbalance. To address this, new data will be incorporated to support model training and refinement.

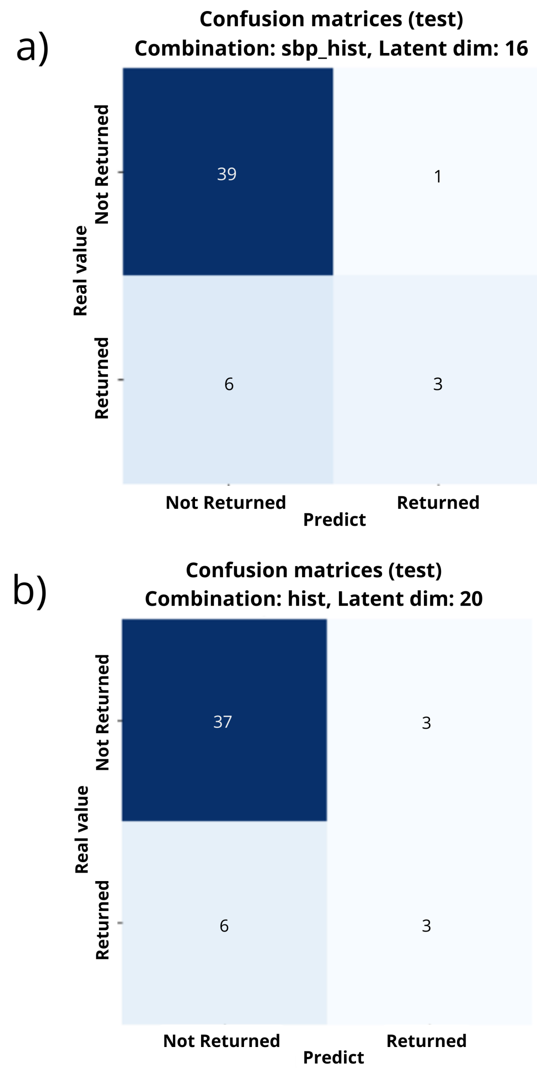

Figure 3. Confusion matrices from the test set using all three folds concatenated. a) SBP + history with latent dimension 16. b) history with latent dimension 20

Table 2. Average results for different combinations of input variables

Combination	Dimension	Mean F1-Score	Balanced Accuracy	ROC-AUC
SBP + History	16	$83.41\% \pm 3.22$	$65.38\% \pm 2.22$	$71.48\% \pm 10.86$
History	20	$80.13\% \pm 5.35$	$62.82\% \pm 3.84$	$59.27\% \pm 8.63$
SBP + DBP + HR-ECG + History	16	$77.04\% \pm 6.84$	$61.53\% \pm 10.01$	$71.48\% \pm 6.26$
DBP + History	8	$81.11\% \pm 7.17$	$61.11\% \pm 9.62$	$70.87\% \pm 7.85$
SBP + DBP + History	20	$81.11\% \pm 7.17$	$61.11\% \pm 9.62$	$68.31\% \pm 3.89$

Moreover, there is a need for improved understanding of the most influential features via feature importance analysis. In this context, AI explainability techniques will be applied once the expanded dataset is available. Thus far, features have been evaluated by testing all possible input combinations (SBP, DBP, HR-ECG, HR-PPG, SpO_2, and History). This approach has identified the most influential variables, with patient history consistently demonstrating the strongest impact on performance. With the inclusion of new data, the application of feature importance methods is expected to further clarify the model's decision-making process and contribute to the development of more interpretable and reliable AI tools for clinical monitoring.

4. Conclusions

The results of the proposed model is encouraging and have achieved high accuracy (≈83%) even with fewer inputs (SBP + history). Through different combinations of input variables, patient history demonstrated strong performance in classification tasks, yielding favorable results across multiple evaluation metrics. However, improvements are still necessary, particularly for cases involving readmitted patients. This can be addressed through the inclusion of additional data to improve predictive reliability, as well as optimization of the model architecture, ultimately supporting its future use as a clinical monitoring tool.

Acknowledgments

The data used in this work were collected by the team from the Instituto do Coração de São Paulo (InCor-SP). The results presented in this work were funded by Samsung Eletrônica da Amazônia Ltda., under the terms of the Brazilian Informatics Law 8.248/91. The research was approved by IRB-CAAE 80963624.2.0000.0068.

References

[1] Nazarian S, Lam K, Darzi A, Ashrafian H. Diagnostic Accuracy of Smartwatches for the Detection of Cardiac Arrhythmia: Systematic Review and Meta-analysis. Journal of Medical Internet Research August 2021;23(8):e28974.

[2] Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Hosseinzadeh S, Sajadi Tabar SS. Smartwatches in healthcare medicine: assistance and monitoring; a scoping review. BMC Medical Informatics and Decision Making November 2023;23(1):248. ISSN 1472-6947.

[3] Chen Z, Yeo CK, Lee BS, Lau CT. Autoencoder-based network anomaly detection. In 2018 Wireless Telecommunications Symposium (WTS). IEEE, April 2018; 1–5.

[4] Monteiro R, Rabello GCM, Moreno CR, Moitinho MS, Pires FA, Samesina N, César LAM, Tarasoutchi F, Fernandes F, Martins PCCO, Mariano BM, Soeiro ADM, Palhares A, Pastore CA, Jatene FB. Enhancing cardiac postoperative care: a smartwatch-integrated remote telemonitoring platform for health screening with ECG analysis. Frontiers in Cardiovascular Medicine September 2024;11.

Address for correspondence:

Júlia Amancio Ferreira
Recod.ai Lab - Av. Albert Einstein, 1251 - Cidade Universitária, Campinas - SP, Brazil, 13083-852
julia.amancio.ferreira10@gmail.com