



# Article A Machine Learning Based Discharge Prediction of Cardiovascular Diseases Patients in Intensive Care Units

Kaouter Karboub <sup>1,2,3,\*</sup> and Mohamed Tabaa <sup>4,\*</sup>

- <sup>1</sup> FRDISI, Hassan II University Casablanca, Casablanca 20000, Morocco
- <sup>2</sup> LRI-EAS, ENSEM, Hassan II University Casablanca, Casablanca 20000, Morocco
- <sup>3</sup> LGIPM, Lorraine University, 57000 Metz, France
- <sup>4</sup> LPRI, EMSI, Casablanca 23300, Morocco
- \* Correspondence: kaouter.karboub@gmail.com (K.K.); m.tabaa@emsi.ma (M.T.); Tel.: +212-661-943-174 (M.T.)

Abstract: This paper targets a major challenge of how to effectively allocate medical resources in intensive care units (ICUs). We trained multiple regression models using the Medical Information Mart for Intensive Care III (MIMIC III) database recorded in the period between 2001 and 2012. The training and validation dataset included pneumonia, sepsis, congestive heart failure, hypotension, chest pain, coronary artery disease, fever, respiratory failure, acute coronary syndrome, shortness of breath, seizure and transient ischemic attack, and aortic stenosis patients' recorded data. Then we tested the models on the unseen data of patients diagnosed with coronary artery disease, congestive heart failure or acute coronary syndrome. We included the admission characteristics, clinical prescriptions, physiological measurements, and discharge characteristics of those patients. We assessed the models' performance using mean residuals and running times as metrics. We ran multiple experiments to study the data partition's impact on the learning phase. The total running time of our best-evaluated model is 123,450.9 mS. The best model gives an average accuracy of 98%, highlighting the location of discharge, initial diagnosis, location of admission, drug therapy, length of stay and internal transfers as the most influencing patterns to decide a patient's readiness for discharge.

**Keywords:** cardiovascular diseases; discharge; Electronic Health Records; intensive care units; machine learning

# 1. Introduction

Factors such as blood pressure, high cholesterol levels and the adoption of bad habits including smoking and highly fat-saturated foods led to double the number of patients with cardiovascular diseases in the period between 2000 and 2019 compared to 1990, according to the American Hospital Association's (AHA) report published in 2020 [1]. In 2017, the Kaiser Family Foundation Analysis of the Organization for Economic Co-operation and Development (OECD) [2], reported that the United States of America (USA), compared to other developed countries, has fewer medical resources (2.6 practicing physicians and 2.8 beds per 1000 population compared to 5.2 and 7.4 in Austria, 4.3 and 8 in Germany, per 1000 population, respectively). On the other hand, data published by the AHA in 2018 [3] indicates there are a total of 5256 registered community hospitals in the United States of which 2704 (more than 51%) deliver intensive care services with a total of 96,596 Intensive Care Unit (ICU) beds. In total, 68,558 of these beds are dedicated to adults (46,795 medicalsurgical, 14,445 for cardiac care and 7318 for other ICU needs), 22,901 for neonatal care, and 5137 are pediatric ICU beds. Geographically, the distribution is mainly in metropolitan areas with 74% of ICU beds followed by 17% in micro-Politian areas and 9% in rural areas. Such disparities arise from a lack of study, biased data, or a misunderstanding of the healthcare ecosystem. In its 2018 report, the Health Systems for Prosperity and Solidary by the World Health Organization (WHO) [4] mentioned financial crisis, political choices, variations in epidemiology and social preferences or variations in efficiency as the main



Citation: Karboub, K.; Tabaa, M. A Machine Learning Based Discharge Prediction of Cardiovascular Diseases Patients in Intensive Care Units. *Healthcare* 2022, *10*, 966. https://doi.org/10.3390/ healthcare10060966

Academic Editor: Tin-Chih Toly Chen

Received: 14 April 2022 Accepted: 9 May 2022 Published: 24 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). reasons why some countries would not invest in the healthcare system. As for the uncertain nature of hospitals' ecosystems, static planning of these medical resources seems to be an inconvenient solution. In the literature, many studies have been attracted by the complexity of such issues. Thus, most of these studies put on the surface the importance of dynamic predictions when trying to be one step ahead of ecosystem changes [5–8].

In the course of solving resource allocation problems in uncertain environments, such as ICUs, researchers focused on two points: (1) the huge possibility of using increasingly leveraged clinical data captured from Electronic Health Records (EHR) systems. (2) The need to predict patient outcomes as a step toward an efficient decision-making tool.

In fact, severity score systems were developed to predict a patient's outcome and to compare quality-of-care and stratification for clinical trials [9]. The development of scoring systems involves a complex combination of clinical acumen and advanced statistical techniques. These scoring systems must be rigorously assessed in terms of accuracy, reliability, and methodological rigor before being introduced into clinical practice [10]. Most of them were first derived from a database of patients and their various physiological measurements during their ICU stay. Used databases are retrospectively analyzed to find which of the selected variables are the most predictive for a chosen outcome. The testing of a model on such a validation cohort cannot be considered to represent independent validation. The sample size and randomization process used to select this cohort from the starting database makes it inevitable that the development and validation samples perform interchangeably. However, in an upgraded special article, a group of internationally recognized clinical experts suggested that severity of illness scores should not be considered as a transition condition from ICU to lower acuity care wards. Instead, they specified that such prioritization identifiers might be used to assess high-risk populations after discharge and not their readiness to be transferred [11]. In fact, a review of multiple studies revealed that cardiovascular patient health status was evaluated from four main angles: (1) their cardiovascular disease history, the demographic and socio-economic factors [12–14], etc. (2) Their health status before and after medical interventions [15,16]. (3) The main psychological and behavioral factors interfering with their medical condition [17–19]. (4) Using the patient's health status as a predictor for future outcomes such as mortality [20,21]. Taking these factors combined, it seems that even studies that focus on scoring system types, case studies, and guidelines for using such risk assessment tools [22-24] can come to the conclusion that severity score systems cannot be used to predict patients' discharge.

On the other hand, a new trend in learning methods or machine learning developed in and for technology and healthcare industries offers tremendous potential to enhance medical research and clinical care, especially as providers increasingly employ EHR.

There are many areas that can benefit from the application of machine learning techniques in the medical field, mainly diagnosis and outcome prediction.

In fact, many studies, such as [25], showed that MIMIC-III specific machine learning models using only 10 clinical variables outperformed nine commonly used severity scoring methods. However, other related studies [26,27] have also shown that machine learning models outperform severity scores in predicting in-hospital mortality. Furthermore, developing health system specific prediction models using machine learning enables continuous improvements of the model by including more training data (as more data becomes available), adding new clinical or laboratory variables to the model, or re-training the model using newly-developed machine learning algorithms.

Following this axis and trying to explore new ways to use machine learning other than simulation and modeling in medicine, we developed multiple machine learning models to predict patients' readiness for discharge who are admitted to ICUs. In other words, when a patient is admitted to ICU, they can be transferred to a lower care unit or sent home after receiving the necessary treatment and staying in the hospital for a period. The goal here is to predict the length of stay in one and only one ward in ICU, any other transfer is considered a discharge. In fact, accurate discharge prediction will enable decision-makers (healthcare providers in most cases) to have a clearer vision of future actions and prevent subsequent readmissions. We made sure to eliminate redundant elements and normalized measures, as the MIMIC-III is collected using CareVue and MetaVision clinical information systems. Our approach, based on correctly imputed missing elements in our dataset, uses different algorithms to compare the performance of different configurations of these models. In fact, these models will predict patients that are more likely to be discharged among other patients. The performance of our model is assessed using a Root Mean Square Error (residual mean) with respect to related characteristics: selection of variables and weights, used variables (age, origin, chronic health status, physiology, and acute diagnosis) and size of the validation population. Such a study is meant to identify the impact of drug therapy, type and time of admission, and processed transfers from non-ICU wards and an ICU's cardiology department on the patient's readiness for discharge. Moreover, to discover new opportunities using a variety of machine learning techniques other than previously mentioned machine learning models, or discrete event simulations that have studied resource allocation problems [28], to provide the percentage of patients that meet the discharge criteria and features learned in the model's training process. This might help physicians to prioritize inspections—to assess discharge—to specific patients. The deployment of such models will aid in the decision-making process of healthcare workers by improving the prediction of premature deaths, making medical decisions about highrisk patients more efficient, evaluating the effectiveness of new treatments, and detecting changes in clinical practices. The rest of the paper is organized into two main sections. The first represents a literature review of main studies in the same scope, followed by a context and mathematical description of patients' flow, then by the implementation, results, and discussion of these results compared to the available state of the art literature. We end the paper with a conclusion.

### 2. Literature Review

Discharge planning is, by consensus, suffering from a lot of variability in the clinical decision-making processes. Most ICUs do not use written patient discharge guidelines. Clinicians have rather little secure evidence upon which to base any decision about discharge location. Such ambiguity can lead to poor management of patients, which can result in premature discharge and, subsequently, death or readmission. This has been a factor in the motivation to create critical care outreach teams and triage models to improve discharge outcomes.

In fact, many studies have tried to determine the factors that can predict the length of stay, or discharge in general, where patients are limited to a certain health condition or group of health conditions [29,30]. We found many models with varying degrees of data specification and accuracy. In the paper [31], hospital records were analyzed to determine if any factors could predict hospital Length Of Stay (LOS) and readmission after colorectal resection through linear regression. The data used in this study is a combination of databases from the National Cancer Registry (NCR) and the Hospital In-Patient Enquiry Scheme (HIPE), which contains records of patients in Ireland. STATA (statistical software for data science) was used to determine the best variables for logistic regression using a combination of likelihood ratio tests. For the LOS, it was determined that age, higher levels of co-morbidities, and marital status were associated with an increased LOS. Another study analyzed hospital records to identify the predictors of an increased LOS after Acute Exacerbation of Chronic Obstructive Pulmonary Disease (AECOP) [32]. A multivariate logistic regression model was created to assess the predictors of early discharge in a period longer than 11 days. The results from the aforementioned study show that being admitted between Thursday and Saturday, having high PaCO2, low serum albumin level, or having heart failure, diabetes, or stroke are the most important predictors of a very long LOS. The LOS of patients with cardiac problems was the focus of [33], which is one of the few that employs machine learning techniques. The patient data was retrieved over a five-year period from a hospital in Iran that specializes in treating and researching cardiovascular conditions. Thirty-six different attributes were included per row and three different models were run on the same data: decision tree, neural network, and support vector machine. Out of the three, the support vector machine approach outperformed the two other models, with the diagnosis ICD-9 (International Statistical Classification of Diseases ninth version) code (this provides an internationally standardized code per disease), the diastolic blood pressure (blood pressure in the arteries between heartbeats) and age being the three most prominent input variables (highest relative weight). In the paper [34], a retrospective review of a database was conducted to determine the predictive factors for hospital stay and mortality. An analysis was performed on a database from the Cleveland Clinic who had undergone noncardiac surgery within a five-year period along with measurements of Mean Arterial Pressure (MAP), Bispectrality Index (BIS), and Minimum Alveolar Concentration (MAC). Through logistic regression, it was found that a "triple-low" value of MAP, BIS, and MAC were strongly correlated with an extended LOS.

In fact, simulation has been extensively used to evaluate the impact of resource availability and the organizational settings on healthcare outcomes quality and the costs related to medical interventions and patients' stay [35,36]. There are several methods of simulation but most commonly they are classified into four main categories: Monte Carlo [37,38], Discrete Event Simulation [39,40], System Dynamics [41,42], and Agent Based Simulation [43,44]. Some of the very important advantages of using these simulation techniques are: 1. The ability to perform "what if" analysis that evaluates the performance of the system in different scenarios, considering many types of input data and model parameters as well as identifying critical points related to the system's bottlenecks [45]. While 2. is the ability to perform these scenarios in different time windows [46]. As a result, simulation is useful when a problem exhibits significant uncertainties that require stochastic analysis.

Indeed, several studies have been conducted on the application of simulation as an effective tool to improve processes in healthcare systems to minimize healthcare costs and increase the satisfaction of patients [47].

The main problems in the healthcare system that are addressed for emergency patients based on simulation and modeling knowledge are resource allocation, and patient flow problems, while for non-elective patients it is mostly scheduling and bed assignment [48–51]. Many studies have been conducted to optimize processes and patient flow in the healthcare systems [52–56]. The optimized patient flow is defined as a high patient throughput, low patient waiting times and short LOSs, while keeping staff utilization rates high and reducing staff idle time. The increasing cost of providing high-quality health care has made hospital administrators minimize resources while still striving to provide the service with the desirable quality. Many studies [57] find simulation modeling attractive since it can estimate the operational characteristics of a complex system as well as monitoring the results of changes in planning and resource allocation prior to implementation, which minimizes the financial risks for decision makers. According to the field of study of this thesis, the literature review of the Discrete Event Simulation (DES) in validation-simulation studies is classified into two categories including patient flow and resource allocation.

However, it does not seem possible to elaborate on a universal, conclusive procedure for matching the most suitable simulation technique to a specific problem.

On the other hand, there are relatively few studies on discharge prediction in a critical care setting: they are exclusively focused on discharge readiness [58,59] or they are designed to predict a specific discharge destination [60]. For example, [61] used demographic, ICU admission, and ICU clinical data measured during the first 24 h of ICU admission to develop a predictive algorithm for the early identification of ICU patients with a high probability of discharge to a long-term acute care hospital. The study found that their predictive algorithm can accurately predict the likelihood of patients' discharges. In addition, [62] investigated the relationship between vitamin D status at ICU admission and the home or non-home discharge destination for critically ill surgical patients. They suggested that vitamin D levels may impact patient-oriented outcomes in ICU, and it might be a modifiable risk factor for the discharge destination.

Alternatively, the prediction of patients' discharge from ICU can be expanded to focus on the characteristics of the patients admitted—many tools have been developed to support discharge planning. Mainly, these tools try to predict the likelihood of complications during hospitalization. Furthermore, these tools try to predict functional adverse outcomes, which can pose serious difficulties during the discharge process [63]. Most of these tools are appropriate for patients admitted for medical conditions, and the majority of those are condition-independent and can be widely applied.

For all these reasons, more data are needed about those factors already present before proper intervention and that is associated with a longer LOS or a discharge with the need for additional care, and that can be investigated in the early phases of the treatment trajectory. The aim of this study was to investigate those factors in a large sample of patients, in order to better understand what can be done to predict, as early as possible, which patients will need personalized and more demanding discharge planning, and possibly to suggest general items suitable for this prediction in general care departments for patients with cardiovascular diseases.

## 3. Context and Methods

The determination of appropriate medical resource distribution in healthcare facilities is a very challenging task as it must be coordinated at three different levels [64]. At the macro-allocation level, patterns of this distribution are drawn by legislation, government funding mandates and healthcare insurance plans. At the organizational level, policies, clinical practice guidelines and protocols decide how resources might be allocated to make maximum use of limited resources. However, at a micro-allocation level, it is a physician's mission to decide whether a treatment or an investigation is in a patient's best interest or not [65,66].

#### 3.1. Mathematical Context

In every patient's discharge, the bed's occupancy distribution is evaluated. Having more visibility on when and how to discharge patients give hospitals' policymakers and physicians in ICUs more flexibility and the ability to draw admission patterns, face admission peaks and manage general wards and lower medical care units more effectively. Moreover, and more importantly, to deliver medical care to the maximum number of patients by reducing the LOS and increasing admission rates. Figure 1 represents the generic guiding flow map divided into two main units: admission flows and the ICUs' In and Out flows. The model presented in Figure 1 has been evaluated and validated using Non-Homogenous Discrete Time Markovian Chains [67].

In ordinary circumstances, a patient in a critical condition might be admitted directly from the emergency room, be a planned admission, or from internal wards in which he or she was admitted as a non-critical case, then needed intensive medical care.

We have developed a model that aims to provide a generic representation of a hospital's internal flows [68]. The aforementioned is based on discrete Markov chains and is validated using real-world data. The following is, therefore, intended to prove how it is important to optimize the number of inspections using machine learning techniques. This goes along with shortening the LOS but also to take into consideration the patient's condition. In a more illustrative image, the mathematical model we are representing here will guide the usage of our dataset. This means that, at every step of the flow, we will take parameters that might impact the patient's readiness for discharge at the end of their treatment.



Figure 1. Flow dynamic in ICUs.

pathology the number admitted patients of Let  $\chi_{k,j,t}(t)$ be of  $k = \{ coronary artery disease, congestive heart failure, acute coronary syndrome \} admitted$ as  $j = \{Emergency, Elective, internal ward\}$  department after spending t time in the hospital. O<sub>k,i,t</sub> represents the overflowing patients from both the external and internal admissions of previously named pathologies. This overflow is calculated considering the patients in hospital and still not served, and scheduled patients not arriving in time. In our model, we consider that all causes leading to overflow can result in very long diagnosis times.

At a given time window, we consider  $W\{1...q\}$  as the number of ICUs wards. We define the bed occupancy function as representing if a ward q is occupied by a patient of pathology or condition type k.

$$U_{q,k} = \frac{number \ of \ patients \ of \ pathology \ k \ allocated \ to \ ward \ q}{total \ number \ of \ beds \ in \ ward \ q}$$
(1)

Let  $F_k = \omega_k - \sum_k U_{q,k}$  be the number of free beds in ward q. The  $D_{q,k}(.)$  reflects the bed's distribution in a given time interval and  $\omega_k$  is the number of beds in ward k.

$$D_{q,k} = \left( \begin{pmatrix} U_{11} & \cdots & U_{1M} \\ \vdots & \ddots & \vdots \\ U_{M1} & \cdots & U_{MM} \end{pmatrix}, (F_1, F_2, \dots, F_M) \right)$$
(2)

where  $(U_{11}, U_{22}, ..., U_{MM} | k = q)$  is the number of patients with primary hospitalizations.

We also define  $\alpha_k$  and  $\beta_{q,k}$  as the primary and secondary hospitalizations' rate represented in Equations (3) and (4), respectively:

$$\alpha_k \equiv \alpha_{q,k} = \frac{\sum_{q,k(k=q)} U_{q,k}}{\sum_{q,k} D_{q,k}}$$
(3)

$$\beta_{q,k} = \frac{\sum D_{q,k} - \sum U_{q,k} - \sum_q F_q}{\sum_{q,k} D_{q,k}}$$
(4)

Based on the work presented in [43], we can conclude that the distribution of the newly arriving patient to the different wards follows the process shown below:

$$\begin{cases} \chi_{k,j,t} * \alpha_k \text{ if } F_k > 0 \text{ for } k \in M \\ \chi_{k,j,t} * \beta_{q,k} \text{ if } F_k = 0 \text{ and } F_{q \neq k} > 0 \text{ for } q \in M \end{cases}$$
(5)

In the matrix above, a newly arriving patient is allocated to the preferred ward in cases where there exists a free bed. Otherwise, the patient is oriented to another ward that may serve a similar service as shown in Figure 1.

We define in the following, that the time spent by a patient of type k as an inpatient in ward q is denoted  $\tau_{q,k}^k$ .

We also define  $Q_{q,k}^n$  as the number of patients of type k in ward q right after the  $n^{th}$  inspection. We can assume that the inspection and discharge patterns can be described by the same distribution and are the same from an operational perspective. The dynamics of such distribution can be cited as follows:

$$\begin{cases} Q_{q,k}^{n+1}(\alpha_{k}=1) = Q_{q,k}^{n} + (1 - \sum_{k \in M} \beta_{q,k}) * (\chi_{k,j,t})^{n+1}) - \xi_{q}^{n+1} for \ every \ k \\ Q_{q,k}^{n+1}(\alpha_{k}<1) = Q_{q,k}^{n} + (1 - \alpha_{k}) * (\chi_{k,j,t})^{n+1} - \xi_{k}^{n+1} \end{cases}$$
(6)

where  $\xi_a^{n+1}$  is the number of patients discharged in the (n + 1)th inspection and

$$\Gamma^{N} = \sum_{n=0, \forall q, \forall k}^{N} \tilde{\xi}_{q}^{n+1}$$
(7)

In the same way, a decision to discharge a patient takes into account the patient's pathology, condition and the period already spent as an inpatient. The main goal of expressing the number of inspections per period is to relate this factor with the service rate and occupancy function. Thus, in used data, we will have to divide the 24 h into specific intervals and set a mean number of inspections. This will help derive a discharge rate per period, which will provide visibility to how many primary and secondary hospitalizations are made and how many patients to admit.

#### 3.2. Data Description

In our study, we used the MIMIC-III, which contains de-identified health-related data of more than 40,000 patients with more than 50,000 hospitalizations in the ICUs of the Beth Israel Deaconess Medical Center in Boston, Massachusetts, in the period between 2001 and 2012. The data in the MIMIC-III is divided into 26 tables, each one comprises a specific type and flow of data such as demographics, vital signs measurements, laboratory test results, procedures, medications, caregiver notes, imaging reports and discharge mortality.

The tables are linked by primary identifiers such as subject-ID, hadm-ID and ICUStay-ID. The recorded measurements are provided by Philips and iMDSoft tools used in CareVue and MetaVision clinical information systems, respectively [69].

The de-identification of patients was incorporated in accordance with the Health Insurance Portability and Accountability Act (HIPAA) standards and the federal code of the USA. It included the removal of the eighteen identification data elements defined by the HIPAA such as patient name, phone number, address, and dates such as the date of birth, date of admission, etc. In the following, we calculated the age of the patients and used their associated dates in an intervals-based approach by which dates are shifted into the future, between 2100 and 2200, in a manner to preserve the intervals [70,71].

For reasons of having the highest mortality rates [72,73] and being responsible for the utilization of a significant proportion of healthcare resources [74], we managed to solely use data related to adult patients with specific types of pathologies as shown in Figure 2. The current study basically includes 4402 admissions and a total of 4226 patients, as some patients are admitted more than once with different admission identifiers. In total, 2804 of

them were admitted to the emergency room, 1466 were elective patients, and 132 were admitted as urgent. Of the total number, 2808 were diagnosed with coronary artery disease, 1315 with congestive heart failure, and 279 with acute coronary syndrome while admitted to ICUs and before receiving any prior treatments. Table 1, below, resumes the patients' related data, input characteristics, outcome characteristics and actual measurements performed.



Figure 2. MIMIC III dataset generation process.

Table 1. Baseline patients' characteristics and outcome measures.

|                          | Overall Population<br>Characteristics | Dead at Discharge<br>Characteristics | Alive at Discharge<br>Characteristics |
|--------------------------|---------------------------------------|--------------------------------------|---------------------------------------|
| Average age              | 65                                    | 73                                   | 64                                    |
| Sex (% men)              | 53%                                   | 54%                                  | 57%                                   |
| Admission type:          |                                       |                                      |                                       |
| Emergency                | 2804                                  | 226                                  | 2577                                  |
| Elective                 | 1466                                  | 14                                   | 1451                                  |
| Urgent                   | 132                                   | 11                                   | 121                                   |
| Type of ICU:             |                                       |                                      |                                       |
| Coronary artery disease  | 2808                                  | 54                                   | 2753                                  |
| Congestive heart failure | 1315                                  | 175                                  | 1140                                  |
| Acute coronary syndrome  | 279                                   | 22                                   | 257                                   |

|  | Overall Population<br>Characteristics | Dead at Discharge<br>Characteristics | Alive at Discharge<br>Characteristics |
|--|---------------------------------------|--------------------------------------|---------------------------------------|
| Average heart rate (bpm)                                   | 80                                    | 100–110                              | 60–100                                |
| Average respiratory rate (cpm)                             | 21                                    | 12–20                                | $\leq 12 \text{ or } \geq 20$         |
| Prescription drugs:<br>Cholesterol lowering<br>medications | 12.83%                                | 0.14%                                | 99.86%                                |
| ACE inhibitors   | 14.62%                                | 0.38%                                | 99.62%                                |
| Bronchodilators  | 10.66%                                | 0.47%                                | 99.53%                                |
| Diuretics  | 9.1%                                  | 1%                                   | 99%                                   |
| Insulins   | 7.85%                                 | 1.04%                                | 98.96%                                |
| Anticoagulants   | 7.42%                                 | 0.8%                                 | 99.2%                                 |
| Electrolytes   | 13.22%                                | 0.66%                                | 99.34%                                |
| Beta blockers  | 7.2%                                  | 1.78%                                | 98.22%                                |
| Antiplatelet agents and DAPT                               | 3.46%                                 | 3.58%                                | 96.42%                                |
| Anti-histamines  | 3.44%                                 | 0.53%                                | 99.47%                                |
| Quinolone antibiotics                                      | 3.2%                                  | 1.42%                                | 98.58%                                |
| Nitrates   | 2.63%                                 | 1.21%                                | 98.79%                                |
| Peptides   | 1.36%                                 | 1.01%                                | 98.99%                                |
| Glucose elevating agents                                   | 1.63%                                 | 5.88%                                | 94.12%                                |
| Antidysrhythmics   | 0.89%                                 | 4.6%                                 | 95.4%                                 |
| Calcium channel blockers                                   | 0.35%                                 | 3.95%                                | 96.05%                                |
| Sulfonic acid  | 0.15%                                 | 6.06%                                | 93.94%                                |
| Hospital length of stay                                    | 2.9 days                              | 3.1 days                             | 2.7 days                              |

# Table 1. Cont.

## 3.3. Preprocessing

Coronary artery disease is a condition where the coronary arteries are narrowed or blocked causing chest pain, congestive heart failure or/and an acute coronary syndrome condition. Such a blockage in the blood supply to the heart muscle area, or ischemia, can cause heart tissues to die within a few minutes. To relieve pain by reducing the heart's workload, and to prevent chest pain and acute coronary symptoms from happening, doctors usually use nitrates, beta blockers, calcium channel blockers, ACE inhibitors, statins, antiplatelet and anticoagulants drugs, respectively (and sometimes combined) to reverse coronary artery narrowing or to open a blocked artery. Although, on the one hand, heart failure is a disorder related to the heart's inability to follow the body's demands in terms of blood flow, congestion of blood and regular beating; this condition is often caused by cardiac causes such as coronary artery disease, myocarditis, heart valve disorders, as well as some non-cardiac causes such as high blood pressure, anemia, kidney failure and others. The treatment of such a condition may, thus, include diuretics and nitrates to relieve symptoms related to the pain such as angiotensin, ACE inhibitors, beta blockers and aldosterone antagonists to help the treatment to succeed.

In the process of understanding the physiological parameters of patients, we noticed that some of these parameters are recorded at a lower frequency compared to others, such as non-invasive blood pressure. Such missing data we could handle using invasive blood pressure, which was continually monitored. In other cases, missing data were replaced by the median of the concerned variable such as the LOS. We also extracted informative patterns from patients' physiological data and the chronological order of events between admission and discharge. Those features were first identified from MIMIC-III admission, chart events, ICU stays and prescription tables based on their known logistic and clinical relevance to our target endpoints. We determined the univariate importance of each feature with respect to the target variable. Such a technique enabled us to make our dataset easier to interpret and to understand how data is distributed within this specific population of patients as it addresses the interrelations between the different timing and physiological

features and discharge decisions. Figure 3 represents the feature selection based correlation of chronological features related to the time when patients were admitted and discharged from ICUs, in addition to the importance of the features related to prescriptions and physiological patterns.



(b)

Figure 3. Cont.





**Figure 3.** Feature selection of informative features. (a,b) represent the correlation between drug dosage and type, and time of discharge, respectively. Figure (c) is the time that marks the end of a specified drug therapy. (d,e) are the chronological timing of admission to ICUs, and their actual discharge time.

As shown in Figure 3a,b, the use of metoprolol, vancomycin, 0.9% sodium chloride, insulin, heparin, 5% dextrose (noting that it is also abbreviated as D5W in the MIMIC-III database), iso-osmotic dextrose, ondansetron ODT, phenytoin, piperacillin-Tazobactam-NA

are used to prevent heart attacks by lowering blood pressure, preventing the formation of blood clots, and fighting bacterial infections. Although a constant rate of admissions to the emergency department and surgery is more likely to be followed by a peak in discharges, as might be noticed, admission and discharge from ICUs follow the patterns of global admission—the more admissions increased, the more ICU admissions and discharges increased and decreased, respectively (Figure 3c,d).

On the other hand, the starting and ending dates are more likely to be constant, which means that patients of similar pathology take similar durations under a given description. Drugs such as cholesterol lowering medications, insulins, ACE inhibitors, diuretics, beta blockers, glucose elevating agents and calcium channel blockers and their combinations designed, in most cases, for arrhythmia, artery disease, coronary disease, chest pain and hypotension have the highest patients' early discharge probability. In the dataset files, there is no indication of an undesirable event experienced by a patient related to drug therapy that interferes with achieving the desired goal of such therapy.

#### 4. Results and Implementation

Severity of illness is a composite of the magnitude of the acute disease, the patient's physiological reserve, and the concurrent level of treatment and organ system support. Of these three variables, the physiological reserve is the most difficult to quantify and modify.

It is generally assessed using functional capacity, co-morbid disease, and age. Loss of functional capacity is an important predictor of frequent hospitalizations and death, and co-morbid disease impacts ICU and hospital outcomes [75].

In our study, feature importance reflects how useful and valuable every single attribute was when building the decision to discharge a patient. This importance is calculated for each attribute in our dataset, by which all attributes are ranked and compared to each other. Every single importance index is calculated by the average training loss reduction gained when using a feature for splitting. As noticed in Figure 3, the discharge location, ICU admission related diagnosis, admission location, drug, out-time (at which patients are transferred from the ICU unit to a recovery ward or to home), frequency of the given drugs, ICU stay ID (which reflects if a patient has an ICU admission history), start date (the date on which drug therapy started), LOS (length of stay), end date (the day in which the drug therapy ended), subject ID (which reflects, if associated with different admission ID, a patient's readmission), the admission date, and the times by which the patient is admitted to ICU are the most informative features. As a consequence, we are using these features as input for our models. Missing values of numerical data are imputed using a basic algorithm of decision trees, while categorical data are encoded and linearly regularized. We experimented with eight models to identify the most suitable one. We have also set the decision tree's regression as the baseline to estimate performance. Further, three of these models: 1. AVG blender; 2. Advanced Generalized Linear Regression Model; and 3. Efficient Neural Network, are ensembles of other models. Due to their iterative nature, Gradient Boosted models are almost guaranteed to overfit the training data, given enough iterations. Table 2 represents the tuning parameters of the baseline models, while all the other models are sophisticated forms of them. Any other parameters not mentioned in the table are left in their default values.

Table 2. Parameters of models.

| Decision Trees        | Gradient Boosted Models    |
|-----------------------|----------------------------|
| Criterion: Entropy    | Number of estimators: 2000 |
| Max depth: 10         | Learning rate: 0.3         |
| Splitter: Best        | Criterion: MSE             |
| Max features: log2    | Min sample leaf: 2         |
| Min samples leaf: 4   | Min samples split: 5       |
| Min samples split: 10 | * *                        |



|                       |                        |                           | - ·             | _        |
|-----------------------|------------------------|---------------------------|-----------------|----------|
| Elevena / dataila the | deve a sector of the o | different al a guitle mag | and in array as |          |
| FIGHTE 4 GETAILS THE  | avnamics of the        | autherent algorithms      | insea in air ai | nnroach  |
| i iguit i utuito ute  | aynamico or the        | unicicili argoritimo      | uscu m our u    | pprouen. |

|   | Auto-tuned<br>stochastic Gradient<br>Descent Regression              | - | Decision Tree<br>regression  | <br>Extreme Gradient<br>Boosted Trees<br>Regression with Early<br>Stopping | 1 | Gradient Boosted<br>Greedy Trees<br>Regression with<br>Early Stopping |
|---|--|---|--|--|---|---|
| Å | Light Gradient<br>Boosted Trees<br>Regression with<br>early stopping |   | Advanced<br>Generalized Linear<br>Regression (Gaussian<br>distribution)<br>Fig.4. Architecture of<br>the proposed approach | <br>Efficient Neural<br>Network (ENET)                                     |   | AVG Blender   |

Figure 4. Architecture of the proposed approach.

The overall workflow used for building the predictive models in the present work is illustrated in Figure 4. We have used Python 3.9 as a programming language to build our models. We used Pandas [76], NumPy [77], SciPy [78], TensorFlow [79], and time [80]. To ensure that our results are not biased towards a specific learning algorithm and to minimize the risk of over-fitting, we have implemented a set of machine learning predictors. To achieve the best performance, it was necessary to find the best hyper-parameters for each algorithm. A grid-search was conducted to find the best parameters for each model. A grid-search combines all possible combinations in a parameter grid where one defines the possible values for each hyper-parameter. In other words, it provides an exhaustive method to evaluate the combinations of all hyper-parameters.

The best possible parameter settings found using grid-search for each algorithm are presented in Table 2:

These methods were used to predict the discharges in a week. To study the impact of the data partition on the models' performance, we conducted four experiments on different data partition configurations. The configurations and results after simulations are

**Table 3.** Performance comparison summary between RM values, accuracy, and prediction time metrics.

|       |                   |                              | Linear<br>Regression |         | Trees Re | gression |           | Ν         | lixed (Blender | s)       |
|-------|-------------------|------------------------------|----------------------|---------|----------|----------|-----------|-----------|----------------|----------|
|       | Ν                 | ſodel                        | LR *                 | TR1 *   | TR2 *    | TR3 *    | TR4 *     | B1 *      | B2 *           | B3 *     |
|       | Acc               | Accuracy %                   |                      | 0.783   | 0.917    | 0.761    | 0.726     | 0.98      | 0.935          | 0.961    |
|       |                   | Validation<br>(18%)          | 0.00682              | 0.18    | 0.00399  | 0.01714  | 0.01353   | -         | -              | -        |
| Exp 1 | RM                | Validation<br>(36%)          | 0.00575              | 0.09    | 0.00261  | 0.00953  | 0.01357   | -         | -              | -        |
|       |                   | Holdout<br>(72%)             | 0.00539              | -       | 0.00173  | 0.00521  | 0.00312   | 0.0012    | 0.00158        | 0.00058  |
|       | Predictio         | on time (mS)                 | 5627.83              | 3734.79 | 6914.43  | 22,443.8 | 17,452.56 | 123,450.9 | 123,800        | 69,376.8 |
|       | Acc               | uracy %                      | 0.88                 | 0.77    | 0.91     | 0.77     | 0.756     | 0.978     | 0.91           | 0.956    |
|       |                   | Validation<br>(18%)<br>Cross | 0.00679              | 0.18    | 0.00399  | 0.01714  | 0.01353   | -         | -              | -        |
| Exp 2 | RM                | Validation<br>(36%)          | 0.00571              | 0.09    | 0.00261  | 0.00953  | 0.01357   | -         | -              | -        |
|       |                   | Holdout                      | 0.00512              | -       | 0.00159  | 0.00503  | 0.00298   | 0.0012    | 0.00147        | 0.00054  |
|       | Predictio         | on time (mS)                 | 5628.01              | 3699.45 | 6999     | 23,166.4 | 20,514.6  | 123,551   | 132,500        | 70,015.2 |
|       | Acc               | uracy %                      | 0.89                 | 0.78    | 0.914    | 0.76     | 0.754     | 0.937     | 0.914          | 0.915    |
|       |                   | Validation<br>(18%)          | 0.00659              | 0.18    | 0.00373  | 0.01694  | 0.0112    | -         | -              | -        |
| Exp 3 | RM                | Validation<br>(46%)          | 0.00551              | 0.085   | 0.00191  | 0.00958  | 0.01057   | -         | -              | -        |
|       |                   | Holdout<br>(72%)             | 0.00512              | -       | 0.00159  | 0.00503  | 0.00298   | 0.80319   | 0.00147        | 0.00054  |
|       | Predic            | ction (mS)                   | 5568                 | 4697    | 7859     | 23,894.2 | 25,735.02 | 215,693   | 151,236        | 78,020   |
|       | Acc               | uracy %                      | 0.89                 | 0.781   | 0.925    | 0.709    | 0.781     | 0.95      | 0.965          | 0.89     |
|       |                   | Validation<br>(25%)          | 0.00614              | 0.1     | 0.00329  | 0.01658  | 0.01123   | -         | -              | -        |
| Exp 4 | RM                | Validation<br>(36%)          | 0.00541              | 0.09    | 0.00261  | 0.00953  | 0.01357   | -         | -              | -        |
|       |                   | Holdout<br>(72%)             | 0.00481              | -       | 0.00148  | 0.00493  | 0.00298   | 0.000001  | 0.9947         | 0.00054  |
|       | Run tii<br>Predic | me for 100<br>tions (mS)     | 5750                 | 3610.9  | 6990     | 32,548   | 24,590.4  | 133,511   | 15,039.3       | 699,081  |

summarized in Table 3.

\* LR: Auto-tuned stochastic Gradient Descent Regression \* TR1: Decision Tree; \* TR2: Extreme Gradient Boosted Trees Regression with Early Stopping \* TR3: Gradient Boosted Greedy Trees Regression with Early Stopping \* TR4: Light Gradient Boosted Trees Regressor with Early Stopping. \* B1: Advanced Generalized Linear Regression Model (GLRM); \* B2: Efficient Neural Network (ENET); \* B3: AVG Blender.

The model achieved an optimized training accuracy value of 0.98 and a mean residual = 0.0004 when all selected features and an advanced generalized linear regression model were used. In Figure 5, below, the discharge location, diagnosis, admission location, drug, out-time (day of the month), prod-strength, day (in a month) in which drug therapy starts, length of stay, (day in the month) in which drug therapy ends, admission time, admitted for the first time or not, time (in the day) in which a patient is admitted to an ICU, and the time in which the patient is supposed to be discharged to a general ward are the most relevant features, with 99.97%, 88%, 74%, 72%, 68%, 65%, 40%, 39%, 38%, 30%, 28%, 28% and 23% of the weights associated to all attributes in accordance with their importance in the training phase, respectively. This approach works by monitoring the performance of our models and automatically selecting the inflection point where performance on the test dataset starts to decrease while performance on the training dataset continues to improve.



Figure 5. MIMIC III: training results.

As mentioned previously, the figure above represents spots or prediction areas within which a given discharge readiness estimation is given. The spots are estimated using the Regressor Fit algorithm. Patients readmitted to emergency within 6 months of their first admission, patients with respiratory failure, and patients with low rates of multi-vitamins are the patients more likely to stay longer in hospital. In addition to this, the discharge location—home, or to a long-term care facility—is a key predictor in deciding if a patient can be discharged or not. Figure 6 represents the mean residuals obtained for each model.

In general, mixed or ensemble algorithms outperformed the other models. Firstly, in terms of performance, where advanced generalized linear regression, efficient neural network, and AVG blenders showed an accuracy of more than 90%. Secondly, in terms of robustness, where in multiple data partitions, these models also showed great results. The recapitulation of these results is presented in Table 3.

It can be seen in Figure 5 that by delaying the prediction somewhat into the admission time, a better prediction can be made. This result is in line with the hypothesis that complementing the data from the emergency service with the data collected at the ward would increase performance. Important features, such as suspected diagnosis and planned transfer to other wards, are added at this stage and this type of specific information can be valuable, especially when combined with all of the lab parameter values. For the admission stage, the balanced probability to be discharged ranges from 0.6 to 0.62. These results were in line with the ones obtained in previous LOS studies. The fact that the patient group used in this dataset was quite broad shows that there is a potential for these machine learning algorithms to be used in more generalized settings in hospitals. Patients using certain types of drugs, such as Hydralazine are more likely to be discharged after 5 days of admission. Earlier research has usually very specific limitations on patient type, such as diabetes or brain surgery patients, while this project focused on a certain hospital clinic and considers interference from outside factors. One of the limiting factors was low precision for the long-staying class (i.e., patient contacts staying longer than 3 days). At the discharge stage, this value ranged from 0.4 to 0.48. Such low precision would make it troublesome to use the prediction in a real system.



**Figure 6.** Residuals of different models (yellow represent expected outputs and blue ground-truth output).

## 5. Discussion

In general, most discharge guideline reports published by the WHO [81], the Society of Critical Care Medicine (SCCM) [82] or the AHA [83], published between 2016 and 2018, list the following discharge criteria for ICU patients: stable hemodynamic parameters, stable respiratory status and airway patency, oxygen requirements not more than 60%, intravenous inotropic support and vasodilators are no longer necessary, cardiac dysrhythmias are controlled, neurologic stability with control of seizures, patients who require chronic mechanical ventilation resolved, and patients with tracheostomies who no longer require frequent suctioning. Discharge planning is multi-factorial and a succession of consecutive parts. One of these parts is the patient's readiness for discharge. The assessment of such treatment transition enables a provider to estimate the patient's and family's ability to leave the original medical care institution and to move to a home health phase or to a lower intensity care area. A readiness for discharge from an ICU assessment requires the evaluation of the patient's physiological stability, cognitive and psychomotor ability to carry out self-management regimens, social support availability and permanent access to healthcare systems [84]. Assessing the patient's readiness for discharge from an ICU is a necessary task for the patient's care and the equitable usage of the ICU's available resources. In both cases, for ICU-Emergency admitted patients or ICU-Elective patients (postoperative monitoring and medical interventions), the determination of discharge line between intensive care and recovery care may withdraw occupancy rates of ICU beds. Discharge criteria enable a triage-based decision; allowing patients to leave ICUs if not necessarily needed, and by increase the rate of readmission.

Both admission and discharge involve a change of location with the potential for gaps in communication that may result in diminished and discontinuous care. There is growing research evidence showing that the outcomes of intensive care are affected by the timing of admission and discharge decisions, which in turn, are influenced by resource availability in the ICU and probable inexpert care on the ordinary wards [85]. Admission to the ICU from 00:00–07:00 h, and at weekends is associated with a higher

mortality, as is discharge from the ICU to the ordinary ward at night [86]. Readmission to intensive care is associated with a hospital death rate 2–10 times that of non-readmitted patients [87] and can be mitigated by intensive care outreach in the form of intensivist-led rapid response teams [88]. Of the high-risk surgical patients admitted to intensive care in 28 European countries, 43% of deaths occurred after discharge to the ordinary ward [89], which suggests there are potential opportunities to look at again with regards to the way in which a patient's readiness for discharge is assessed. Unintentional discontinuation of chronic medications is also common following discharge from the ICU and is associated with adverse patient outcomes. Decisions to admit patients to ICU or to discharge them to the ward are determined by the severity of their illness.

Our approach can be used in two different pathways. Firstly, to identify patients who are most likely to be discharged in each day per week and give them an uncertainty rate. Following these recommendations, hospital staff can prioritize these patients to be inspected, and then discharge them as early as possible so that other patients can be admitted to ICUs. Secondly, the possibility of ranking patients into mild and moderate severity while discharged. This allows the hospital to prioritize who and when to run the remaining tasks for patients on the discharge list. Despite the many advantages of our approach, it is important to highlight that the performance of our models may improve with a larger dataset in terms of accuracy but may exponentially decrease in terms of execution time.

A range of different tools and methods have previously been proposed, as shown in the comparative Table 4, with the aim of improving ICU discharge practice. These tools range from criteria for evaluating discharge readiness [90,91], to guidelines for discharge planning and education [92].

Predictive models based on very large patient numbers capture more population information than the individual clinician can acquire in a lifetime; however, the clinician will know more about the individual patient than any of these models can. For this reason, predictive systems may inform clinical judgement, but cannot replace it. Triage protocols to maximize the use of scarce resources in high seasonality periods have been modelled prospectively and retrospectively [93], demonstrating theoretical value in releasing intensive care beds by denying admission to those categorized as being too well or too sick to benefit. Several models have been developed to inform safe and timely ICU discharge decisions. Simple univariate risk factors include prolonged LOS, unstable vital signs (including tachypnea or tachycardia) and poor pulmonary function. For example, [94] have modelled post-ICU mortality and ICU readmission using data from more than 700,000 patients, incorporating admission diagnosis, severity of illness, laboratory values, and physiological variables in the last 24 h of the ICU stay. The Stability and Workload Index for Transfer score [95], and a model developed in France [96], have similar predictive precision for ICU readmission. Others have identified the potential for important reductions in mortality had triage models been used to avoid premature ICU discharge.

Table 4. Results and Benchmarking.

| Ref  | Methods and Approach   | Dataset   | Metrics and Results  | Scoring of Recommendation<br>Strength        |
|------|--|---|--|--|
| [97] | Focus: prognostication of<br>clinical outcomes in ICUs.<br>Methods: multivariate<br>imputation by chained<br>equations for missing data<br>imputation. Adaboost, parRF<br>(parallel implementation of a<br>random forest),<br>SVMRadialWeights (SVM with<br>radial basis function kernel and<br>class weights), avNNet<br>(averaged Neural Network) and<br>deep NN as classifiers. | Critical Care Health<br>Informatics Collaborative<br>(CCHIC) data infrastructure<br>(22,514 intensive care<br>admissions of which 21,911<br>were used in the study; 90.8%<br>of them were alive at<br>discharge.) | On day 2 (AUC):<br>parRF: 0.853 simple and 0.857 cumulative.<br>avNNet: 0.864 simple and 0.879 cumulative.<br>Adaboost: 0.862 simple and 0.879<br>cumulative. svmRadialWeights: 0.849<br>simple and 0.884 cumulative. DeepNN:<br>0.881 simple and 0.895 cumulative | Larger data improves model's<br>performance. |

| Ref       | Methods and Approach   | Dataset   | Metrics and Results   | Scoring of Recommendation<br>Strength                 |
|-----------|--|---|---|---|
| [98]      | Focus: mortality prediction,<br>LOS prediction and ICD-9 code<br>group prediction.<br>Methods: SAPS-II (Simplified<br>Acute Physiology Score) and<br>SOFA. Super learner models,<br>RNN and FFN.   | Medical Information Mart for<br>Intensive Care III (MIMIC-III)<br>(v1.4)  | SuperLearner-I: AUROC = 0.8448 and<br>AUPRC = 0.4351.<br>SuperLearner-II: AUROC = 0.8701 and<br>AUPRC = 0.4991.<br>FFN: AUROC = 0.8496 and<br>AUPRC = 0.4632.<br>RNN: AUROC = 0.8544 and<br>AUPRC = 0.4519.<br>MMDL: AUROC = 0.8664 and<br>AUPRC = 0.4776.<br>Scoring methods: 0.8035 AND 0.7322 for<br>AUROC (SAPS-II and SOFA respectively),<br>AUPRC: 0.3586 for SAPS-II and 0.3191 for<br>SOFA.   | Larger data improves model's performance.             |
| [99]      | Focus: prediction of final<br>diagnosis and clinical outcomes.<br>Methods: universal language<br>model fine-tuning for text<br>classification (ULMFiT)   | Medical Information Mart for<br>Intensive Care III (MIMIC-III)  | Accuracy: 80.3% for diagnosis top10, 80.5%<br>procedure top10, 70.7% diagnosis top50,<br>63.9% procedures top50.  | Larger data improves model's performance.             |
| [100]     | Focus: in-hospital mortality<br>prediction. Methods: deep<br>learning networks.  | Medical Information Mart for<br>Intensive Care III (MIMIC-III)<br>(42,818 hospital admissions of<br>35,348 patients)  | Mortality prediction: AUROC: 0.9178 with<br>data of all sources (AS) and 0.9029 with<br>chart data (CD). PRAUC: 0.6251 for AS,<br>0.5701 for CD.<br>LOS prediction: AUROC: 0.8806 for AS and<br>0.8642 for CD. PRAUC of 0.6821 and 0.6575,<br>respectively with AS and CD.  | Larger data improves model's performance.             |
| [101]     | Focus: ICUs discharge<br>prediction. Methods: random<br>forest (RF) and logistic classifier<br>(LC).   | Bristol Royal Infirmary<br>general intensive care unit<br>(GICU) (1870 intensive care<br>patients) and 7592 from<br>MIMIC-III.  | On the MIMIC dataset: AUROC(RF):0.8859,<br>AUROC(LC): 0.8726. Accuracy (RF): 0.8531<br>and accuracy (LC): 0.8494. sensitivity (RF):<br>0.9049 and sensitivity (LC) is 0.9001.   | Larger data improves model's performance.             |
| [102]     | Focus: prediction of discharge<br>location in ICUs. Methods:<br>National Early Warning Score<br>(NEWS/NEWS 2)  | Surgical, coronary, cardiac<br>surgery recovery, medical and<br>trauma surgical intensive care<br>patients with single<br>admission in ICUs in a US<br>hospital.  | The NEWS AUROC (95% CI): all patients<br>0.727 (0.709–0.745); Coronary Care Unit<br>(CCU) 0.829 (0.821–0.837); Cardiac Surgery<br>Recovery Unit (CSRU) 0.844 (0.838–0.850);<br>Medical Intensive Care Unit (MICU) 0.778<br>(0.767–0.791); Surgical Intensive Care Unit<br>(SICU) 0.775 (0.762–0.788); Trauma Surgical<br>Intensive Care Unit (TSICU) 0.765<br>(0.751–0.773).  | Larger data improves model's performance.             |
| [103]     | Focus: risk scoring in ICUs.<br>Methods: attentive deep<br>Markov model (AttDMM).  | MIMIC-III with 53,423 ICU stays.  | AttDMM with AUROC of 0.876.   | Not specified   |
| [104]     | Focus: ICU readmission<br>prediction after 24 to 72 h of<br>discharge.<br>Methods: fuzzy modeling and<br>tree search feature selection<br>technique.   | MIMIC II (data of 4 different<br>ICUs 26,655 patients, of<br>which 19,075 are adults; 38%<br>of the adult patients stayed at<br>the medical ICU (MICU), 27%<br>at the surgical ICU (SICU),<br>20% at the cardiac surgery<br>recovery unit (CSRU) and<br>15% at the critical care unit<br>(CCU)) | AUROC of 0.76 and <i>p</i> -value of 0.006 with<br>sequential forward selection. AUROC of<br>0.68 and <i>p</i> -value under 0.05 with sequential<br>backward elimination.   | Not specified   |
| [105]     | Focus: length of stay prediction<br>in ICUs. Methods: neural<br>network (random forest)  | MIMIC-III (31,018 chosen data points)   | Accuracy of 80%.  | Larger dataset might improve the model's performance. |
| Our model | Preprocessing: Regularized<br>Linear Processing, Ordinal<br>encoding of categorical<br>variables, Tree based Algorithm.<br>Prediction tools: Auto-tuned<br>stochastic Gradient Descent<br>Regression, Decision Tree,<br>Extreme Gradient Boosted Trees<br>Regression with Early Stopping,<br>Gradient Boosted Greedy Trees<br>Regression with Early Stopping,<br>Light Gradient Boosted Trees<br>Regression with Early Stopping,<br>Advanced Generalized Linear<br>Regression Model (GLRM),<br>Efficient Neural Network<br>(ENET), AVG Blender | MIMIC-III   | <ul> <li>{18%<sup>1</sup>, 36%<sup>2</sup>, 72%<sup>3</sup>}: Accuracy = 98%,<br/>RESIDUAL MEAN = 0.000001, Prediction<br/>time: 123,450.9 mS.</li> <li>{18%<sup>1</sup>, 36%<sup>2</sup>, and 80%<sup>3</sup>}: Accuracy: 97.8%,<br/>RESIDUAL MEAN = 0.000001, Prediction<br/>time: 123,551mS.</li> <li>{18%<sup>1</sup>, 46%<sup>2</sup>, 72%<sup>3</sup>}: Accuracy: 93.7%,<br/>RESIDUAL MEAN = 0.000319, Prediction<br/>time: 215,693 mS.</li> <li>{25%<sup>1</sup>, 36%<sup>2</sup>, 72%<sup>3</sup>}: Accuracy = 96.5%,<br/>RESIDUAL MEAN = 0.00147 and<br/>Prediction time: 15,039.3 mS.</li> <li><sup>1</sup>: Validation, <sup>2</sup>: Cross validation, <sup>3</sup>: Holdout</li> </ul> | Larger dataset improve<br>model's performance         |

# Table 4. Cont.

### 6. Conclusions and Perspectives

To optimize hospitals' resources and maximize service quality, we presented a method to handle CVD patients' discharge from ICUs using their physiological data and history of treatments, and on the learning capacity of ML algorithms. Firstly, we included physiological, hospital internal transfers associated with patients and prescriptions in their treatments from the MIMIC-III ICUs database recorded in the period between 2001 and 2012. The database contains de-identified, health-related data of more than 40,000 patients with more than 50,000 hospitalizations. The used dataset includes pneumonia, sepsis, congestive heart failure, hypotension, chest pain, coronary artery disease, fever, respiratory failure, acute coronary syndrome, shortness of breath, seizure, transient ischemic attack, and aortic stenosis that have the same ICU admission, flow mutation and prescription patterns. Multiple ML algorithms were used to compare the effectiveness of each model. We ran multiple experiences on the processed dataset to study the importance of the dataset's size in every learning phase of our models. We tested our models using data related to 4226 cardiovascular disease patients. As a result, we achieved a better accuracy performance of 0.98 and an RM (Residual Mean) = 0.0004 using advanced generalized linear models, which included stochastic and trees regression. These results encourage future work that will include studying the impact of such a decision support tool on internal logistics and post-discharge outcomes. All methods need prospective validation.

Furthermore, the dataset available in this project contained a lot of information that was not included in the training and testing data due to the amount of feature engineering it would have required. As the whole pipeline, from raw data to polished flat-table format, had to be implemented, not all possible information could be extracted due to time limitations. This is one of the most interesting points to explore in similar projects in the future. The dataset contained copious amounts of time-series data from different lab tests, and it would be very interesting to develop more features connected to trends present in these types of data. Only statical features were used in this project, but if features could be extracted that reflect how, for example, a vital parameter has varied over time, this could prove very valuable to the ML models as trends are very important when clinicians evaluate patients. It would, therefore, be interesting to explore the feature engineering aspects of this dataset more extensively, not only for the time-series data.

Author Contributions: All authors have contributed equally and homogeneously to the achievement of this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by the Moroccan School of Engineering Sciences EMSI Casablanca and Foundation of Research and Development and Innovation of Engineering Sciences FRDISI.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: MIMIC III Dataset.

Conflicts of Interest: The authors declare no conflict of interest.

## References

- Virani, S.S.; Alonso, A.; Benjamin, E.J.; Bittencourt, M.S.; Callaway, C.W.; Carson, A.P.; Chamberlain, A.M.; Chang, A.R.; Cheng, S.; Delling, F.N.; et al. Heart Disease and Stroke Statistics—2020 Update: A Report from the American Heart Association. *Circulation* 2020, 141, e139–e596. [CrossRef] [PubMed]
- Organisation for Economic Co-Operation and Development (OECD). Health Care Resources. Available online: https://stats.oecd. org/# (accessed on 12 February 2022).
- Bouch, M.C.F.E.D.C.; Thompson, J. Severity scoring systems in the critically ill. *Contin. Educ. Anaesth. Crit. Care Pain* 2008, 8, 181–185. [CrossRef]
- 4. Mitton, C.; Smith, N.; Peacock, S.; Evoy, B.; Abelson, J. Public participation in health care priority setting: A scoping review. *Health Policy* **2009**, *91*, 219–228. [CrossRef] [PubMed]
- 5. Makridakis, S.; Kirkham, R.; Wakefield, A.; Papadaki, M.; Kirkham, J.; Long, L. Forecasting, uncertainty and risk; perspectives on clinical decision-making in preventive and curative medicine. *Int. J. Forecast.* **2019**, *35*, 659–666. [CrossRef]

- 6. Han, P.K.J.; Klein, W.M.P.; Arora, N.K. Varieties of Uncertainty in Health Care: A Conceptual Taxonomy. *Med. Decis. Mak.* 2011, 31, 828–838. [CrossRef]
- Luo, L.; Luo, L.; Zhang, X.; He, X. Hospital daily outpatient visits forecasting using a combinatorial model based on ARIMA and SES models. *BMC Health Serv. Res.* 2017, 17, 469. [CrossRef]
- 8. Dhawale, T.; Steuten, L.M.; Deeg, H.J. Uncertainty of Physicians and Patients in Medical Decision Making. *Biol. Blood Marrow Transplant.* 2017, 23, 865–869. [CrossRef]
- 9. Guo, T.; Fan, Y.; Chen, M.; Wu, X.; Zhang, L.; He, T.; Wang, H.; Wan, J.; Wang, X.; Lu, Z. Cardiovascular Implications of Fatal Outcomes of Patients With Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol.* **2020**, *5*, 811–818. [CrossRef]
- 10. Mehra, M.R.; Desai, S.S.; Kuy, S.; Henry, T.D.; Patel, A.N. Cardiovascular Disease, Drug Therapy, and Mortality in COVID-19. *N. Engl. J. Med.* **2020**, *382*, e102. [CrossRef]
- 11. Wong, L.E.; Hawkins, J.E.; Langness, S.; Murrell, K.L.; Iris, P.; Sammann, A. Where Are All the Patients? Addressing COVID-19 Fear to Encourage Sick Patients to Seek Emergency Care. *NEJM-Catal.* **2020**, *119*, 187–189. [CrossRef]
- 12. Jamshidi, S.; Parker, J.S.; Hashemi, S. The effects of environmental factors on the patient outcomes in hospital environments: A review of literature. *Front. Arch. Res.* **2019**, *9*, 249–263. [CrossRef]
- Mommersteeg, P.M.; Denollet, J.; Spertus, J.A.; Pedersen, S.S. Health status as a risk factor in cardiovascular disease: A systematic review of current evidence. *Am. Heart J.* 2009, 157, 208–218. [CrossRef] [PubMed]
- Rumsfeld, J.S.; Alexander, K.P.; Goff Jr, D.C.; Graham, M.M.; Ho, P.M.; Masoudi, F.A.; Moser, D.K.; Roger, V.L.; Slaughter, M.S.; Smolderen, K.G.; et al. Cardiovascular health: The importance of measuring patient-reported health status: A scientific statement from the American Heart Association. *Circulation* 2013, 127, 2233–2249. [CrossRef]
- Spertus, J.A.; Nerella, R.; Kettlekamp, R.; House, J.; Marso, S.; Borkon, A.M.; Rumsfeld, J.S. Risk of Restenosis and Health Status Outcomes for Patients Undergoing Percutaneous Coronary Intervention Versus Coronary Artery Bypass Graft Surgery. *Circulation* 2005, 111, 768–773. [CrossRef] [PubMed]
- De Jonge, P.; Ormel, J.; Brink, R.H.V.D.; Van Melle, J.P.; Spijkerman, T.A.; Kuijper, A.; Van Veldhuisen, D.J.; Berg, M.V.D.; Honig, A.; Crijns, H.J.; et al. Symptom Dimensions of Depression Following Myocardial Infarction and Their Relationship with Somatic Health Status and Cardiovascular Prognosis. *Am. J. Psychiatry* 2006, *163*, 138–144. [CrossRef] [PubMed]
- 17. Mallik, S.; Krumholz, H.M.; Lin, Z.Q.; Kasl, S.V.; Mattera, J.A.; Roumains, S.A.; Vaccarino, V. Patients With Depressive Symptoms Have Lower Health Status Benefits After Coronary Artery Bypass Surgery. *Circulation* **2005**, *111*, 271–277. [CrossRef] [PubMed]
- Morgan, A.L.; Masoudi, F.A.; Havranek, E.P.; Jones, P.G.; Peterson, P.N.; Krumholz, H.M.; Spertus, J.A.; Rumsfeld, J.S.; Cardiovascular Outcomes Research Consortium. Difficulty taking medications, depression, and health status in heart failure patients. *J. Card. Fail.* 2006, 12, 54–60. [CrossRef]
- 19. Suwanno, J.; Petpichetchian, W.; Riegel, B.; Issaramalai, S.A. A model predicting health status of patients with heart failure. *J. Cardiovasc. Nurs.* **2009**, *24*, 118–126. [CrossRef]
- De Jong, M.M.J.; Moser, D.K.; Chung, M.L. Predictors of Health Status for Heart Failure Patients. Prog. Cardiovasc. Nurs. 2005, 20, 155–162. [CrossRef]
- Palazón-Bru, A.; Carbayo-Herencia, J.A.; Vigo, M.I.; Gil-Guillén, V.F. A method to construct a points system to predict cardiovascular disease considering repeated measures of risk factors. *PeerJ* 2016, 4, e1673. [CrossRef]
- 22. Brady, W.; De Souza, K. The HEART score: A guide to its application in the emergency department. *Turk. J. Emerg. Med.* **2018**, *18*, 47–51. [CrossRef] [PubMed]
- Than, M.P.; Flaws, D.F.; Cullen, L.; Deely, J.M. Cardiac Risk Stratification Scoring Systems for Suspected Acute Coronary Syndromes in the Emergency Department. *Curr. Emerg. Hosp. Med. Rep.* 2013, 1, 53–63. [CrossRef]
- Bourdeaux, C.; Ghosh, E.; Atallah, L.; Palanisamy, K.; Patel, P.; Thomas, M.; Gould, T.; Warburton, J.; Rivers, J.; Hadfield, J. Impact of a computerized decision support tool deployed in two intensive care units on acute kidney injury progression and guideline compliance: A prospective observational study. *Crit. Care* 2020, 24, 656. [CrossRef] [PubMed]
- 25. Kong, G.; Lin, K.; Hu, Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 251. [CrossRef]
- Lin, K.; Hu, Y.; Kong, G. Predicting in-hospital mortality of patients with acute kidney injury in the ICU using random forest model. Int. J. Med. Inform. 2019, 125, 55–61. [CrossRef]
- Silva, I.; Moody, G.; Scott, D.J.; Celi, L.A.; Mark, R.G. Predicting In-Hospital Mortality of ICU Patients: The PhysioNet/Computing in Cardiology Challenge 2012. Comput. Cardiol. 2012, 39, 245–248.
- Becerra-Fernández, M.; Herrera, M.M.; Trejos, C.; Romero, O.R. Resources Allocation in Service Planning Using Discrete-Event Simulation. *Ing. Univ.* 2021, 25, 1–23. [CrossRef]
- 29. Gruenberg, D.A.; Shelton, W.; Rose, S.L.; Rutter, A.E.; Socaris, S.; McGee, G. Factors influencing length of stay in the intensive care unit. *Am. J. Crit. Care* 2006, *15*, 502–509. [CrossRef]
- Toptas, M.; Samanci, N.S.; Akkoc, I.; Yucetas, E.; Cebeci, E.; Sen, O.; Can, M.M.; Ozturk, S. Factors Affecting the Length of Stay in the Intensive Care Unit: Our Clinical Experience. *BioMed. Res. Int.* 2018, 2018, 9438046. [CrossRef]
- 31. Kelly, M.; Sharp, L.; Dwane, F.; Kelleher, T.; Comber, H. Factors predicting hospital length-of-stay and readmission after colorectal resection: A population-based study of elective and emergency admissions. *BMC Health Serv. Res.* **2012**, *12*, 77. [CrossRef]
- 32. Wang, Y.; Stavem, K.; Dahl, F.; Humerfelt, S.; Haugen, T. Factors associated with a prolonged length of stay after acute exacerbation of chronic obstructive pulmonary disease (AECOPD). *Int. J. Chronic Obstr. Pulm. Dis.* **2014**, *9*, 99–105. [CrossRef] [PubMed]

- Hachesu, P.R.; Ahmadi, M.; Alizadeh, S.; Sadoughi, F. Use of Data Mining Techniques to Determine and Predict Length of Stay of Cardiac Patients. *Healthc. Inform. Res.* 2013, 19, 121–129. [CrossRef] [PubMed]
- Sessler, D.I.; Sigl, J.C.; Kelley, S.D.; Chamoun, N.G.; Manberg, P.J.; Saager, L.; Kurz, A.; Greenwald, S. Hospital stay and mortality are increased in patients having a "triple low" of low blood pressure, low bi-spectral index, and low minimum alveolar concentration of volatile anesthesia. J. Am. Soc. Anesthesiol. 2012, 116, 1195–1203. [CrossRef] [PubMed]
- 35. Jacobson, S.H.; Hall, S.N.; Swisher, J.R. Discrete-event simulation of health care systems. In *Patient Flow: Reducing Delay in Healthcare Delivery*; Springer: Boston, MA, USA, 2006; pp. 251–252.
- Hamrock, E.; Paige, K.; Parks, J.; Scheulen, J.; Levin, S. Discrete event simulation for healthcare or-ganizations: A tool for decision making. J. Healthc. Manag. 2013, 58, 110–124. [PubMed]
- Brailsford, S.C. System dynamics: What's in it for healthcare simulation modelers. In Proceedings of the 2008 Winter Simulation Conference, Miami, FL, USA, 7–10 December 2008; pp. 1478–1483. [CrossRef]
- Faezipour, M.; Ferreira, S. A System Dynamics Perspective of Patient Satisfaction in Healthcare. *Procedia Comput. Sci.* 2013, 16, 148–156. [CrossRef]
- Zhou, J.; Wang, J.; Wang, J. A simulation engine for stochastic timed petri nets and application to emergency healthcare systems. IEEE/CAA J. Autom. Sin. 2019, 6, 969–980. [CrossRef]
- Young, T.; Eatock, J.; Jahangirian, M.; Naseer, A.; Lilford, R. Three critical challenges for modeling and simulation in healthcare. In Proceedings of the 2009 Winter Simulation Conference (WSC), Austin, TX, USA, 13–16 December 2009; pp. 1823–1830. [CrossRef]
- 41. Wang, L. An agent-based simulation for workflow in Emergency Department. In Proceedings of the 2009 Systems and Information Engineering Design Symposium, Charlottesville, VA, USA, 24 April 2009; pp. 19–23. [CrossRef]
- 42. Liu, Z. Modeling and simulation for healthcare operations management using high performance computing and agent-based model. *J. Comput. Sci. Technol.* **2017**, *17*, 87–88.
- 43. Gaba, D.M. The Future Vision of Simulation in Healthcare. Simul. Health J. Soc. Simul. Health 2007, 2, 126–135. [CrossRef]
- 44. Kumar, A.; Krishnamurthi, R.; Nayyar, A.; Sharma, K.; Grover, V.; Hossain, E. A Novel Smart Healthcare Design, Simulation, and Implementation Using Healthcare 4.0 Processes. *IEEE Access* **2020**, *8*, 118433–118471. [CrossRef]
- 45. Mielczarek, B. Review of modelling approaches for healthcare simulation. Oper. Res. Decis. Vol. 2016, 26, 55–72.
- Mao-Guo, G.; Li-Cheng, J.; Dong-Dong, Y.; Wen-Ping, M. Evolutionary multi-objective optimization algorithms (EMO). J. Softw. 2009, 7, 75–81. [CrossRef]
- 47. Conforti, D.; Guerriero, F.; Guido, R. Optimization models for radiotherapy patient scheduling. 40R 2007, 6, 263–278. [CrossRef]
- Kim, Y.-J.; Yoo, J.-H. The utilization of debriefing for simulation in healthcare: A literature review. Nurse Educ. Pract. 2020, 43, 102698. [CrossRef] [PubMed]
- 49. Fetter, R.B.; Thompson, J.D. Patients' waiting time and doctors' idle time in the outpatient setting. Health Serv. Res. 1966, 1, 66–90.
- 50. Fetter, R.B.; Thompson, J.D. The Simulation of Hospital Systems. Oper. Res. 1965, 13, 689–711. [CrossRef]
- Jun, J.B.; Jacobson, S.H.; Swisher, J.R. Application of discrete-event simulation in health care clinics: A survey. J. Oper. Res. Soc. 1999, 50, 109–123. [CrossRef]
- Jin, X.; Sivakumar, A.I.; Lim, S.Y. A simulation based analysis on reducing patient waiting time for consultation in an outpatient eye clinic. In Proceedings of the 2013 Winter Simulations Conference (WSC), Washington, DC, USA, 8–11 December 2013; pp. 2192–2203. [CrossRef]
- 53. Rising, E.J.; Baron, R.; Averill, B. A Systems Analysis of a University-Health-Service Outpatient Clinic. *Oper. Res.* **1973**, *21*, 1030–1047. [CrossRef]
- 54. Evans, G.W.; Gor, T.B.; Unger, E. A simulation model for evaluating personnel schedules in a hospital emergency department. In Proceedings of the Winter Simulation Conference, Coronado, CA, USA, 8–11 December 1996; pp. 1205–1209. [CrossRef]
- 55. Elbattah, M.; Molloy, O. Coupling Simulation with Machine Learning. In Proceedings of the 2016 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation, New York, NY, USA, 15 May 2016; pp. 47–56. [CrossRef]
- Chen, N.; Xie, X.; Zeng, Z.; Zhong, X.; Brenny-Fitzpatrick, M.; Liegel, B.A.; Zheng, L.; Li, J. Improving Discharge Process at the University of Wisconsin Hospital: A System-Theoretic Method. *IEEE Trans. Autom. Sci. Eng.* 2019, 16, 1732–1749. [CrossRef]
- Chand, S.; Moskowitz, H.; Norris, J.B.; Shade, S.; Willis, D.R. Improving patient flow at an outpa-tient clinic: Study of sources of variability and improvement factors. *Health Care Manag. Sci.* 2009, 12, 325–340. [CrossRef]
- Li, Y. Study of cardiovascular disease prediction model based on random forest in eastern China. Sci. Rep. 2020, 10, 5245. [CrossRef]
- 59. Badawi, O.; Breslow, M.J. Readmissions and Death after ICU Discharge: Development and Validation of Two Predictive Models. *PLoS ONE* **2012**, *7*, e48758. [CrossRef] [PubMed]
- 60. Zimmerman, J.E.; Wagner, D.P.; Draper, E.A.; Knaus, W.A. Improving intensive care unit discharge decisions: Sup-plementing physician judgment with predictions of next day risk for life support. *Crit. Care Med.* **1994**, 22, 1373–1384. [CrossRef] [PubMed]
- Cuadrado, D.; Riano, D.; Gomez, J.; BodI, M.; Sirgo, G.; Esteban, F.; Garcıa, R.; Rodrıguez, A. Pursuing optimal prediction of discharge time in icus with machine learning methods. In Proceedings of the Conference on Artificial Intelligence in Medicine in Europe, Poznan, Poland, 26–29 June 2019; Springer: Cham, Switzerland, 2019; pp. 150–154.
- 62. Szubski, C.R.; Tellez, A.; Klika, A.K.; Xu, M.; Kattan, M.W.; Guzman, J.A.; Barsoum, W.K. Predicting Discharge to a Long-Term Acute Care Hospital After Admission to an Intensive Care Unit. *Am. J. Crit. Care* **2014**, *23*, e46–e53. [CrossRef] [PubMed]

- Muhlestein, W.E.; Akagi, D.S.; Kallos, J.A.; Morone, P.J.; Weaver, K.D.; Thompson, R.C.; Chambless, L.B. Using a guided machine learning ensemble model to predict discharge disposition following meningioma resection. *J. Neurol. Surg. Part B: Skull Base* 2018, 79, 123–130. [CrossRef] [PubMed]
- 64. Brook, K.; Camargo, C.A.; Christopher, K.B.; Quraishi, S.A. Admission vitamin D status is associated with discharge destination in critically ill surgical patients. *Ann. Intensiv. Care* **2015**, *5*, 23. [CrossRef] [PubMed]
- Levin, S.; Barnes, S.; Toerper, M.; Debraine, A.; DeAngelo, A.; Hamrock, E.; Hinson, J.; Hoyer, E.; Dungarani, T.; Howell, E. Machine-learning-based hospital discharge predictions can support multidisciplinary rounds and decrease hospital length-of-stay. *BMJ Innov.* 2020, 7, 414–421. [CrossRef]
- 66. Abad, Z.S.H.; Maslove, D.M.; Lee, J. Predicting Discharge Destination of Critically Ill Patients Using Machine Learning. *IEEE J. Biomed. Health Inform.* 2020, 25, 827–837. [CrossRef]
- 67. Jaotombo, F.; Pauly, V.; Auquier, P.; Orleans, V.; Boucekine, M.; Fond, G.; Ghattas, B.; Boyer, L. Machine-learning prediction of unplanned 30-day rehospitalization using the French hospital medico-administrative database. *Medicine* 2020, 99, e22361. [CrossRef]
- Karboub, K.; Tabaa, M.; Dellagi, S.; Monteiro, F.; Dandache, A.; Moutaouakkil, F. Modeling and Validation of the Hospital's Ambulatory and Inpatients Operations Using a Non-Homogenous Discrete Time Markovian Chains. *IEEE Access* 2021, 9, 103044–103055. [CrossRef]
- 69. Ransom, H.; Olsson, J.M. Allocation of Health Care Resources: Principles for Decision-making. Pediatr. Rev. 2017, 38, 320–329. [CrossRef]
- Mitton, C.; Donaldson, C. Health care priority setting: Principles, practice and challenges. Cost Eff. Resour. Alloc. 2004, 2, 3. [CrossRef] [PubMed]
- 71. American Hospital Association. Fast Facts on U.S. Hospitals, 2020. Chicago, IL: American Hospital Association. Available online: https://www.aha.org/statistics/fast-facts-us-hospitals (accessed on 18 March 2020).
- 72. Cylus, J.; Permanand, G.; Smith, P.C.; World Health Organization. Making the Economic Case for Investing in Health Systems. What Is the Evidence That Health Systems Advance Economic and Fiscal Objectives? Available online: http://www.euro.who. int/\_\_data/assets/pdf\_file/0010/380728/pb-tallinn-01-eng.pdf (accessed on 11 February 2022).
- 73. Johnson, A.; Pollard, T.; Mark, R. MIMIC-III Clinical Database (version 1.4). *PhysioNet.* 2016, 3. [CrossRef]
- 74. Johnson, A.E.W.; Pollard, T.J.; Shen, L.; Lehman, L.-W.H.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.; Celi, L.A.; Mark, R.G. MIMIC-III, a freely accessible critical care database. *Sci. Data* 2016, *3*, 160035. [CrossRef] [PubMed]
- 75. Goldfrad, C.; Rowan, K. Consequences of discharges from intensive care at night. Lancet 2010, 355, 1138–1142. [CrossRef]
- 76. Goldberger, A.; Amaral, L.; Glass, L.; Hausdorff, J.; Ivanov, P.C.; Mark, R.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 2000, 101, e215–e220. [CrossRef] [PubMed]
- 77. Python Documentaries. Available online: https://pandas.pydata.org/ (accessed on 11 February 2022).
- 78. Python Documentaries. Available online: https://numpy.org/ (accessed on 11 February 2022).
- 79. Python Documentaries. Available online: https://scipy.org/ (accessed on 11 February 2022).
- 80. Python Documentaries. Available online: https://www.tensorflow.org/learn (accessed on 11 February 2022).
- 81. Python Documentaries. Available online: https://docs.python.org/3/library/time.html (accessed on 11 February 2022).
- Roth, G.A.; Johnson, C.; Abajobir, A.; Abd-Allah, F.; Abera, S.F.; Abyu, G.; Ahmed, M.; Aksut, B.; Alam, T.; Alam, K.; et al. Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. J. Am. Coll. Cardiol. 2017, 70, 1–25. [CrossRef] [PubMed]
- 83. Manfredini, R.; De Giorgi, A.; Tiseo, R.; Boari, B.; Cappadona, R.; Salmi, R.; Gallerani, M.; Signani, F.; Manfredini, F.; Mikhailidis, D.P.; et al. Marital Status, Cardiovascular Diseases, and Cardiovascular Risk Factors: A Review of the Evidence. *J. Women's Health* **2017**, *26*, 624–632. [CrossRef]
- 84. Lavie, C.J.; Arena, R.; Alpert, M.A.; Milani, R.V.; Ventura, H.O. Management of cardiovascular diseases in patients with obesity. *Nat. Rev. Cardiol.* **2018**, *15*, 45–56. [CrossRef]
- Kerac, M.; Trehan, I.; Weisz, A.; Agapova, S.; Manary, M. Admission and Discharge Criteria for the Management of Severe Acute Malnutrition in Infants Aged under 6 Months. Available online: <a href="https://www.who.int/nutrition/publications/guidelines/updates\_management\_SAM\_infantandchildren\_review8.pdf?ua=1">https://www.who.int/nutrition/publications/guidelines/ updates\_management\_SAM\_infantandchildren\_review8.pdf?ua=1</a> (accessed on 11 February 2022).
- 86. Nates, J.L.; Nunnally, M.; Kleinpell, R.; Blosser, S.; Goldner, J.; Birriel, B.; Fowler, C.S.; Byrum, D.; Miles, W.S.; Bailey, H.; et al. ICU Admission, Discharge, and Triage Guidelines. *Crit. Care Med.* **2016**, *44*, 1553–1602. [CrossRef]
- Yancy, C.W.; Jessup, M.; Bozkurt, B.; Butler, J.; Casey, D.E., Jr.; Colvin, M.M.; Drazner, M.H.; Filippatos, G.S.; Fonarow, G.C.; Givertz, M.M.; et al. 2017 ACC/AHA/HFSA Focused Update of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. J. Am. Coll. Cardiol. 2017, 70, 776–803.
- 88. Titler, M.G.; Pettit, D.M. Discharge readiness assessment. J. Cardiovasc. Nurs. 1995, 9, 64–74. [CrossRef] [PubMed]
- Al-Qahtani, S.; Al-Dorzi, H.M.; Tamim, H.M.; Hussain, S.; Fong, L.; Taher, S.; Al-Knawy, B.A.; Arabi, Y. Impact of an Intensivist-Led Multidisciplinary Extended Rapid Response Team on Hospital-Wide Cardiopulmonary Arrests and Mortality. *Crit. Care Med.* 2013, 41, 506–517. [CrossRef] [PubMed]
- Pearse, R.M.; Moreno, R.P.; Bauer, P.; Pelosi, P.; Metnitz, P.; Spies, C.; Vallet, B.; Vincent, J.L.; Hoeft, A.; Rhodes, A. European Surgical Outcomes Study (EuSOS) group for the Trials groups of the European Society of Intensive Care Medicine and the European Society of Anesthesiology. Mortality after surgery in Europe: A 7 day cohort study. *Lancet* 2012, *380*, 1059–1065. [CrossRef]

- Bell, C.M.; Brener, S.S.; Gunraj, N.; Huo, C.; Bierman, A.S.; Scales, D.C.; Bajcar, J.; Zwarenstein, M.; Urbach, D.R. Association of ICU or hospital admission with unintentional discontinuation of medications for chronic diseases. J. Am. Med. Assoc. 2011, 306, 840–847. [CrossRef]
- 92. Knight, G. Nurse-led discharge from high dependency unit. Nurs. Crit. Care 2003, 8, 56-61. [CrossRef]
- Bakker, J.; Damen, J.; Van Zanten, A.R.H.; Hubben, J.H. Admission and discharge criteria for intensive care departments. *Ned. Tijdschr. Voor Geneeskd.* 2003, 147, 110–115.
- 94. Badawi, O.; Liu, X.; Hassan, E.; Amelung, P.J.; Swami, S. Evaluation of ICU Risk Models Adapted for Use as Continuous Markers of Severity of Illness Throughout the ICU Stay. *Crit. Care Med.* **2018**, *46*, 361–367. [CrossRef]
- Desautels, T.; Das, R.; Calvert, J.; Trivedi, M.; Summers, C.; Wales, D.J.; Ercole, A. Prediction of early unplanned intensive care unit readmission in a UK tertiary care hospital: A crosssectional machine learning approach. *BMJ Open* 2017, 7, e017199. [CrossRef]
- 96. Hosein, F.; Bobrovitz, N.; Berthelot, S.; Zygun, D.; Ghali, W.A.; Stelfox, H.T. A systematic review of tools for predicting severe adverse events following patient discharge from intensive care units. *Crit. Care* **2013**, *17*, R102–R110. [CrossRef]
- 97. Araujo, T.G.; Rieder, M.M.; Kutchak, F.M.; Franco Filho, J.W. Readmissões e óbitos após a alta da UTI: Um desafio da terapia intensiva. *Rev. Bras. Ter. Intensiva* 2013, 25, 32–38. [CrossRef]
- Pollard, T.J.; Johnson, A.E.W.; Raffa, J.D.; Celi, L.A.; Mark, R.G.; Badawi, O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci. Data* 2018, *5*, 180178. [CrossRef] [PubMed]
- Harris, S.; Shi, S.; Brealey, D.; MacCallum, N.S.; Denaxas, S.; Perez-Suarez, D.; Singer, M. Critical Care Health Informatics Collaborative (CCHIC): Data, tools and methods for reproducible research: A multi-centre UK intensive care database. *Int. J. Med. Inform.* 2018, 112, 82–89. [CrossRef] [PubMed]
- Nuthakki, S.; Neela, S.; Gichoya, J.W.; Purkayastha, S. Natural language processing of MIMIC-III clinical notes for identifying diagnosis and procedures with neural networks. arXiv 2019, arXiv:1912.12397.
- 101. Rogers, P.; Wang, D.; Lu, Z. Medical Information Mart for Intensive Care: A Foundation for the Fusion of Artificial Intelligence and Real-World Data. *Front. Artif. Intell.* **2021**, *4*, 76. [CrossRef] [PubMed]
- 102. McWilliams, C.J.; Lawson, D.J.; Santos-Rodriguez, R.; Gilchrist, I.D.; Champneys, A.; Gould, T.H.; Thomas, M.J.; Bourdeaux, C.P. Towards a decision support tool for intensive care discharge: Machine learning algorithm development using electronic healthcare data from MIMIC-III and Bristol, UK. *BMJ Open* 2019, 9, e025925. [CrossRef]
- Zaidi, H.; Bader-El-Den, M.; McNicholas, J. Using the National Early Warning Score (NEWS/NEWS 2) in different Intensive Care Units (ICUs) to predict the discharge location of patients. *BMC Public Health* 2019, 19, 1231. [CrossRef]
- Fialho, A.; Cismondi, F.; Vieira, S.; Reti, S.; Sousa, J.M.C.; Finkelstein, S. Data mining using clinical physiology at discharge to predict ICU readmissions. *Expert Syst. Appl.* 2012, 39, 13158–13165. [CrossRef]
- 105. Gentimis, T.; Alnaser, A.; Durante, A.; Cook, K.; Steele, R. Predicting hospital length of stay using neural networks on mimic iii data. In Proceedings of the 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), Orlando, FL, USA, 6–10 November 2017.