Clustering-Based Optimization for Emergency Patient Bed Assignment Problem

Hela Jedidi[1,2], Hajer Ben-Romdhane[2], Issam Nouaouri[3], Saoussen Krichen[2]

Abstract— **This study addresses the critical challenge of hospital bed assignment during emergencies by introducing a two-stage combining unsupervised learning and optimization. In the first stage, clustering techniques (K-means, GMM, DBSCAN, and HDBSCAN) are employed to automatically categorize patients into different risk levels based on clinical features, eliminating the need for labeled training data. In the second stage, a linear programming model optimizes the assignment of patients to hospital beds by minimizing costs while respecting capacity limits and medical priority constraints. The proposed approach is tested and validated on instances derived from real-world datasets from Tunisian hospitals. Our proposal approach significantly reduces the data preprocessing workload while ensuring effective prioritization of critical cases across both small and large-scale scenarios. Additionally, the optimal number of clusters is determined through silhouette analysis, enhancing the clinical relevance of the patient clutering.**

***K*eywords: Unsupervised learning, Patient clustering, Emergency bed assignment, Linear programming, Resource allocation.**

I. INTRODUCTION

In late 2019, a novel coronavirus (SARS-CoV-2) emerged in Wuhan, China, raising a global crisis that deeply disrupted healthcare systems [12]. This highly contagious disease primarily attacks the respiratory system and spreads rapidly between individuals [10]. Within weeks, cases were reported across 216 countries, and by March 11, 2020, the World Health Organization (WHO) officially declared COVID-19 a deadly disease caused by SARS-CoV-2 [12].

The rapid spread of COVID-19 exposed significant weaknesses in healthcare infrastructures, particularly in Tunisia. Hospitals and medical facilities were unprepared for the sudden surge in patient numbers, resulting in an overwhelming demand for critical resources, especially in hospital beds [14]. Emergency departments (EDs) and intensive care units (ICUs) reached full capacity, forcing healthcare providers to make urgent and often difficult decisions regarding patient prioritization. In this context, the challenge of resource allocation emerged, focusing on assigning patients to beds according to their medical condition to ensure that limited resources could be managed efficiently and equitably during the pandemic.

This work introduces a novel integrated solution that combines clustering and optimization techniques. It begins by segmenting patients into distinct clusters based on similarities in their medical conditions, identifying distinct levels of patient risk ranging from mild to critical cases, before assigning them to a specific bed. This risk stratification step enables healthcare providers to prioritize hospital admission for high-risk patients, particularly when resources such as hospital beds are limited. In the subsequent phase, we formulate an optimization problem that allocates these patient groups to appropriate hospital beds according to their risk levels. We solve the Patient Bed Assignment Problem (PBAP) using an exact approach to guarantee optimal utilization of available resources while respecting patient priorities and hospital constraints.

The present paper is organized as follows. Section 2 presents a review of the relevant literature on this topic. Section 3 provides a detailed description of the problem and its mathematical formulation. Section 4 outlines the proposed solution methodology. The computational results are presented in Section 5 to illustrate our approach. The last section details our concluding remarks.

II. RELATED LITTERATURE

This section reviews the main research in patient severity assessment and hospital bed assignment, especially during pandemics. We start by exploring approaches using supervised learning for patient risk prediction, followed by unsupervised clustering techniques for patient stratification. Finally, we explore optimization models designed to tackle PBAP.

In line with this focus, numerous studies have applied Machine Learning (ML) techniques to predict patient medical conditions, both during pandemics and routine emergency department triage, using clinical features to enhance patient care and optimize bed assignment [4]. In 2020, Yan et al. [5] developed an XGBoost model based on specific clinical features to predict death risk, achieving high precision despite limited data. Similarly, Yao et al. [6] employed Support Vector Machines to classify patients into severe and mild categories using clinical and laboratory characteristics. Other work, Schifer et al. [3] employed supervised and unsupervised learning approaches on labeled and unlabeled data to evaluate and predict the severity and criticality of the contagious state of patients. However, most of these models focus on binary classification (e.g. mortality vs. recovery) [7]; there remains a critical need for a finer segmentation of patients, according to the severity of their health status, to ensure the prioritization and assignment of severe patients in hospital beds. This limitation becomes especially critical

---

[1] Esprit School of Business, Tunis, Tunisia `hela.jedidi@isg.rnu.tn`

[2] LARODEC laboratory, Université de Tunis, Institut Superieur de Gestion de Tunis, Avenue la liberté, bouchoucha, Le bardo, Tunisia

[3] LGI2A laboratory, Université de l'Artois, UR 3926, Bethune, 62400, France

during pandemics, where data patterns are often unclear due to the evolving nature of the disease. In such cases, clustering techniques provide a valuable alternative. Unlike supervised methods, which depend on labeled data, clustering can reveal hidden patterns in unlabeled datasets, grouping patients into meaningful categories based on their clinical features.

Few studies have applied clustering when labeled data is limited, yet these techniques are useful for identifying hidden patterns during emergencies. Sadegh et al. [7] used SOM and K-means to cluster patients based on symptoms and outcomes, revealing that while severe symptoms were more common in fatal cases, they didn't always lead to death; age affected ICU stays, while sex had no impact. Boutazart et al. [8] clustered death, confirmed, and recovered cases using K-means and EM, showing that GMM offered more stable insights than K-means alone. Moreover, Julian et al. [9] applied X-means to four key clinical tests, identifying three severity-based clusters. Similarly, Chaudhary et al. [11] combined PCA with K-means to detect cross-country communities linked to virus spread.

While these studies effectively applied clustering, they did not focus on grouping patients based on the severity of their health condition, a crucial aspect that our research seeks to address, looking for facilitating the assignment of patients to specific beds based on their risk levels, in emergency situations. In this regard, the PBAP has evolved with numerous extensions. In the static model, elective patients are allocated to beds for a specified duration, subject to constraints categorized as either hard (strictly enforced) or soft (flexible), based on their characteristics [1]. Given that the PBAP problem is NP-hard [2], various solution approaches have been investigated, including exact algorithms [13] as well as heuristic methods [14]. Furthermore, some researchers have combined supervised learning and optimization techniques to enhance assignment efficiency and reduce costs [3], [15]. In contrast to previous studies that address patient grouping and hospital bed assignment as separate challenges, our work combines these two problems by proposing a hybrid framework that integrates unsupervised learning for severity-based patient clustering with an exact optimization model for bed allocation. This combined approach ensures that patients are prioritized not only based on their clinical condition but also in accordance with hospital resource constraints, a dimension rarely addressed in the reviewed literature.

III. PROBLEM STATEMENT

In the context of the COVID-19 pandemic, hospitals face the critical challenge of efficiently assigning beds to patients with varying levels of severity, while considering the limited bed capacity. The problem of assigning patients to specific beds, taking into account the urgency of their condition, is referred to as the Hospital PBAP (HPBAP). In this scenario, the objective is to group patients according to their clinical characteristics, which reflect their risk level, and then optimize the assignment of beds accordingly.

Patients are admitted to the hospital upon arrival according to their health condition, which is determined by a set of clinical features such as respiratory distress, oxygen saturation levels, comorbidities, etc. Given the urgency and uncertainty of the situation, clustering methods are used to group patients with similar clinical conditions and risk levels, without knowing the number of groups in advance. The goal is to determine natural groupings of patients based on these features, which then inform the decision on how to assign them to available beds.

Suppose that five COVID-19 patients arrive at a hospital during the pandemic, each presenting different clinical characteristics reflecting varying levels of severity. The hospital faces a critical shortage of resources, with a limited number of available beds corresponding to different care levels, such as ICU beds, respiratory beds, and normal beds. In this context, the number of patient groups is not predefined; instead, a clustering approach is applied to their clinical features to automatically identify natural groupings that reveal distinct risk levels.

After performing the clustering process, patients may be categorized into different severity groups, for example, two patients with high-risk and three patients with moderate-risk, depending on the observed similarities in their clinical profiles. Based on these identified groups, an optimization phase is then conducted to assign patients to the most appropriate beds. The objective of this optimization is to prioritize high-risk patients by assigning them to ICU and respiratory beds, which are critical for their care. Only one moderate-risk patient from three will be assigned to a normal bed, ensuring that critical resources are reserved for those in most urgent need.

The main challenge is that patients' risk levels are unknown and must be identified from clinical data. It is therefore necessary to determine risk groupings without predefined categories and allocate beds accordingly. In our research, we aim to identify patients with high contagiousness, group them into distinct risk categories, and determine the optimal assignment of each patient to an appropriate bed based on their health condition.

IV. COMBINING CLUSTERING METHODS AND OPTIMIZATION FOR HPBAP

This paper proposes to solve the HPBAP by incorporating additional constraints compared to the PBAP. Specifically, patients are assigned to bed types that match their clinical condition. Given the NP-hard nature [2] of the proposed problem, we propose a two-stage solution methodology illustrated in Figure 1.

In the first phase, unlike traditional classification approaches, we adopt an unsupervised clustering strategy to group patients based on their clinical characteristic. This enables more informed and targeted bed assignments, particularly useful in emergency settings. In the second phase, these patient groups are used as input to an exact optimization model based on a branch-and-bound approach [20], which identifies the most cost-effective and clinically appropriate assignments while respecting capacity and prioritization constraints.

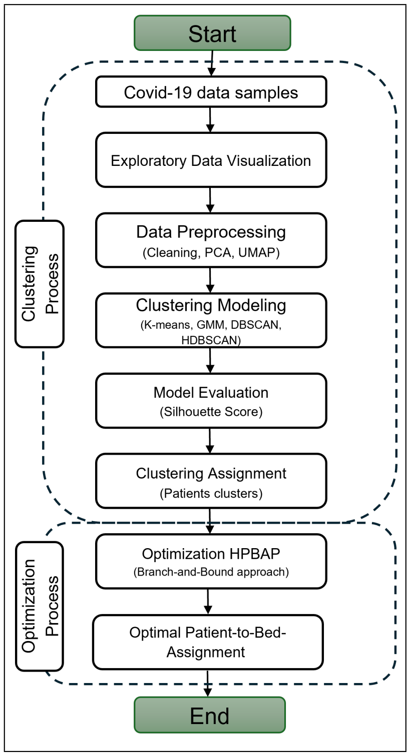

Fig. 1. The proposed methodology flowchart

A. Clustering process

The clustering process provides an effective approach for managing hospital bed availability during emergencies. It helps categorize patients based on their medical profiles and potential contagion, facilitating the allocation of suitable beds. To tackle this challenge, we applied four unsupervised clustering algorithms, aiming to uncover the natural groupings within patient data. Specifically, we evaluate the performance of K-Means, Gaussian Mixture Model (GMM), DBSCAN, and HDBSCAN. These are combined with two dimensionality reduction techniques, Principal Component Analysis (PCA) and Uniform Manifold Approximation and Projection (UMAP), which improve performance and reduce the complexity of high-dimensional data. This approach allows for the identification of meaningful patient clusters without requiring predefined labels. The proposed methodology is validated using a COVID-19 dataset to ensure its applicability in real-world emergency contexts.

1) ***Dataset description****:* The research utilized data obtained from the University Hospital of Sousse, Tunisia, covering September to December 2020. This data set contains demographic and clinical details of patients who tested positive for COVID-19. It comprises 20 instances, divided into Small and Large groups according to different patient numbers and bed availability (refer to Table I). Each set provides details on the number of beds, the planned duration of hospital stays, and the number of patients per category. The different category of instances includes various patient features, as well as: *Age, Glucose (mmol/l), Systolic Blood Pressure (mmHg), Diastolic Blood Pressure (mmHg), Temperature (°C), Heart Rate (beats/minute), Respiratory Rate (breaths/minute), and Oxygen Saturation (%).*

TABLE I

KEY FEATURES OF THE DATASET CATEGORIES (WITH 10 INSTANCES IN EACH CATEGORY)

Categories	Beds	Patients Number	Days
Small set	63	76-253	7
Large set	116	644-1500	7

2) ***Exploratory data visualization****:* Looking for more understanding of the collected data, an exploratory data analysis (EDA) for a small set of data, conducted in this study, provides a first understanding of the dataset's structure. Figure 2 illustrates the distribution of key continuous variables through boxplots, highlighting their spread, medians, and potential outliers.

The analysis shows that continuous variables such as Age, Systolic Blood Pressure, and Temperature have different distribution patterns. For example, age follows an almost normal distribution, while variables such as glucose and respiratory rate are noticeably skewed. Additionally, bivariate analysis reveals a strong positive correlation between Systolic and Diastolic Blood Pressure, suggesting that these two measures tend to increase together. However, most other variable pairs demonstrate weak or no clear linear relationships.

Following the initial data exploration, a preprocessing step is essential to transform and standardize the variables and to apply dimensionality reduction techniques, preparing the data for effective clustering.

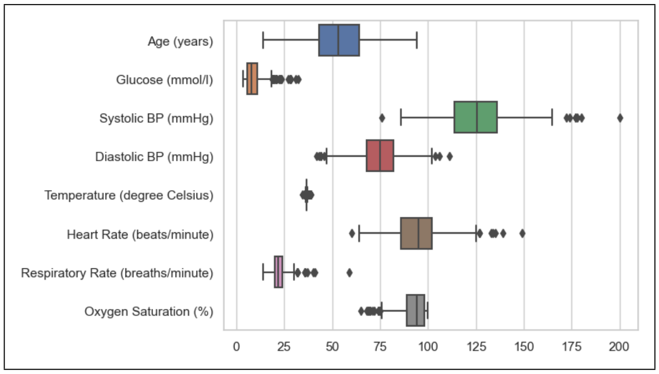

Fig. 2. The distribution of continuous features using Boxplots for a small dataset

3) ***Dataset preprocessing****:* Before applying clustering techniques, a data preprocessing phase was conducted to prepare the dataset for analysis. Since clinical characteristics had different scales, standardization (1) was first used to ensure that each variable contributed equally to the clustering process [16], where z represents the transformed feature value, x is the value of the feature before transformation, μ the average of x and σ the standard deviation of x.

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

Following standardization, Principal Component Analysis (PCA) was applied to decrease the data's dimensionality

while preserving maximum variance. PCA generates new uncorrelated variables, called principal components, that capture the essential structure of the data while minimizing noise [11].
Additionally, Uniform Manifold Approximation and Projection (UMAP) was explored as an alternative dimensionality reduction technique. UMAP is particularly effective in preserving both local and global data structures, making it suitable for datasets like ours, where complex nonlinear relationships between variables might exist [16].

*4) **Clustering modeling**:* In this work, several clustering algorithms were adopted to model the structure of both small and large datasets. The aim was to capture different patterns depending on the nature of the data: linear, skewed, or non-linear. We selected K-Means, Gaussian Mixture Models (GMM), DBSCAN, and HDBSCAN, as they are among the most recognized and widely applied clustering methods in the literature (Xu and Wunsch, 2005). Each technique offers specific advantages depending on the distribution and structure of the data, which we detail below.

- **K-Means** is a clustering technique based on centroids that divides the data into k distinct groups by minimizing within-cluster variance, making it efficient for linearly separable, spherical, and balanced datasets. However, its performance declines with skewed or irregularly shaped data [11].
- **Gaussian Mixture Model (GMM)** is a probabilistic clustering method that represents data as a combination of several Gaussian distributions, allowing flexible cluster shapes and soft membership assignments, making it suitable for overlapping or skewed data [8].
- **DBSCAN (Density-Based Spatial Clustering of Applications with Noise)** is a density-based clustering algorithm that detects clusters of arbitrary shape by grouping closely packed points and marking low-density points as noise, making it effective for non-linear and noisy datasets. Its performance depends heavily on the choice of hyperparameters (epsilon and minimum points) [17].
- **HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise)** extends DBSCAN by building a hierarchical clustering structure based on density and extracting the most stable clusters, improving performance on datasets with varying density and complex non-linear structures [17].

The selection of these algorithms was motivated by the nature of the datasets studied and the need to comprehensively assess performance across different data complexities. K-Means and GMM were primarily applied to well-structured or moderately complex data, while DBSCAN and HDBSCAN demonstrated superior performance in scenarios involving non-linear separations, skewness, or noise. Building upon the insights gained from the clustering phase, we proceed to the next stage of our methodology: optimizing patient allocation decisions through a mathematical modeling approach.

B. Optimization process

The mathematical model introduced in this study extends the HPBAP formulation proposed by [14] by presenting a soft version. Specifically, it relaxes some constraints, permitting controlled violations that are penalized within the objective function. This relaxation increases the model's flexibility in high-demand scenarios such as pandemics, where strict constraint satisfaction may lead to infeasibility or suboptimal utilization of resources. By incorporating penalties into the objective function, the model ensures that any violation remains minimal and is balanced against overall system efficiency. The HPBAP notations are:

- **Sets/Parameters:**
  $p \in P$: Set of patients
  $k \in K$: Set of clusters
  $p \in P_k$: Subset of patients P assigned to cluster k
  $d \in D$: Set of planning days
  $r \in R$: Set of type of beds $R = \{1 : Critical\ \ bed, 2 : Respiratory\ \ bed, 3 : Normal\ \ bed\}$
  D_p : Set of days of patient p hospitalization
  $C_{p,r}$: Cost of assigning patient p to bed r
  Q_r : Total number of beds
  W_{Tr} : Weight of transfer constraint
- **Decision variables** :
  $x_{p,r,d}$: 1 if patient p is assigned to a specific bed r on day d, 0 otherwise
  $t_{p,r,d}$: 1 if patient p is transferred from bed r on day d, 0 otherwise

The mathematical formulation of the HPBAP problem:

- The objective function (2) minimizes two cost components: the cost of assigning patients to inappropriate beds and the cost of transferring patients between beds during their stay.

$$Min \sum_{\substack{p\in P,\\ r\in R, d\in D}} C_{p,r} \times x_{p,r,d} + \sum_{\substack{p\in P,\\ r\in R, d\in D}} W_{Tr} \times t_{r,d} \qquad (2)$$

- Constraints (3) and (4) guarantee that each patient is assigned to only one bed type that is compatible with their identified contagious state (as determined through clustering) at each decision period. Here, $R_k \subset R$ represents the set of bed types allowed for patients belonging to the cluster k.

$$\sum_{r\in R} x_{p,r,d} = 1, \forall p \in P_k, d \in D_p \qquad (3)$$

$$\sum_{r\in R_k} x_{p,r,d} = 1, \forall p \in P_k, d \in D_p \qquad (4)$$

- Constraints (5) ensure that the total number of patients assigned to a specific bed type does not exceed the available capacity for that type at any given time.

$$\sum_{p\in P} x_{p,r,d} \leq Q_r, \forall r \in R, d \in D \qquad (5)$$

- Constraints (6) enforce continuity of bed assignment during the patient's entire stay, penalizing unnecessary transfers between different beds.

$$x_{p,r,d} - x_{p,r,d+1} \leq t_{p,r,d}, \forall p \in P, r \in R, d \in D_p \quad (6)$$

V. EXPERIMENTAL RESULTS

This section presents the detailed results of the experiments conducted according to the previously established methodology. We first evaluated the clustering algorithms (K-means, GMM, DBSCAN, and HDBSCAN) performance without applying any regularization. The silhouette scores obtained were negligible for the small dataset across all methods (0.15, 0.107, 0.17, 0.17), respectively, indicating poor clustering quality. For the larger dataset, only HDBSCAN showed relatively better results (silhouette score 0.33); however, it produced a single cluster, which is not acceptable in our context, where multiple clusters are necessary to correctly assign patients to different beds (figure 3).

To overcome this, we introduced a PCA-based regularization to reduce noise in the datasets, tuning the number of components between 2 and 9, corresponding to the number of features. Following this regularization with 2 components, the best performance was observed with GMM generating three clusters for the large dataset (silhouette score 0.35) and GMM producing two clusters for the small dataset (silhouette score 0.419). These differences can be explained by the varying number of instances and the different patient distributions in each dataset (see figure 4).

To further improve the clustering, we applied UMAP, given that our data contained skewed and non-linear features. We optimized UMAP parameters by varying the number of neighbors (10, 15, 30, 50) and the minimum distance (0.0, 0.1, 0.3, 0.5), identifying the best configuration at 10 neighbors and a minimum distance of 0.0. Under these conditions, K-means performed best for the large dataset by forming 33 clusters with a silhouette score equal to 0.353, while for the small dataset, K-means with two clusters yielded the best results (silhouette score equal to 0.49) after testing different values of k from 2 to 10. Finally, after comparing the different approaches, we concluded that k-means with three clusters after UMAP regularization is the most effective technique for the large dataset with a 0,353 silhouette score (see figure 5), while K-means with two clusters after UMAP regularization provides the best solution for the small dataset with a 0,49 silhouette score. Throughout all experiments, the silhouette score was consistently used as the primary evaluation metric to guide the selection of both clustering algorithms and their corresponding parameters. In the clustering phase, patients were grouped into three clusters using an unsupervised method based on their clinical and demographic features. This data-driven approach revealed that approximately 30–40% of patients were consistently allocated to the first cluster across all datasets. This group was interpreted as representing individuals in a highly contagious or severe condition. These patients typically required intensive medical care and were prioritized for assignment to critical care beds.

The second cluster, comprising around 20–30% of patients, typically represented individuals with moderate symptoms. These patients were expected to need respiratory beds, depending on availability. The third cluster contained the remaining patients, generally classified as mild cases, best suited for normal beds. In cases where respiratory beds were unavailable, patients in the second cluster were redirected to normal beds. If no normal beds were available, mild-risk patients (cluster three) were not hospitalized, following standard triage protocols in resource-constrained environments. The clustering step thus enabled a flexible, adaptive categorization of patients based on actual characteristics rather than a rigid, predefined assignment to bed types. This flexibility is especially important in emergency contexts where patient profiles vary widely and bed resources are limited.

Moreover, the methodology was designed to remain robust under different numbers of clusters. In scenarios where only two clusters are produced, the assignment is straightforward: the first cluster is mapped to critical beds, while the second is mapped to respiratory beds. However, when the algorithm identifies more than three clusters, a merging strategy is applied: the two clusters with the highest severity profiles are combined and collectively treated as priority cases for assignment to critical beds. The subsequent clusters are grouped accordingly for assignment to respiratory and then normal beds, based on their relative risk levels.

To assess the practical implications of the clustering phase, we integrated these clusters into the patient-to-bed assignment model and solved it using IBM ILOG CPLEX Optimization Studio (version 12.10). Table II summarizes the results of the optimization step. The first column ("% Critical patient (cluster)") reports the proportion of patients classified in the most severe cluster (Cluster 1). The second column ("% Critical Admission") shows the percentage of these high-priority patients who were successfully assigned to critical beds. The column "COST" reflects the total assignment cost incurred, and the column "CPU-Time (s)" indicates the computation time of the solver.

As shown in Table II, the proposed clustering-based method outperformed the baseline model (same optimization algorithm executed without the clustering phase) across all scenarios. For the small dataset, clustering enabled 100% admission of critical patients with a reduced cost (7220 vs.

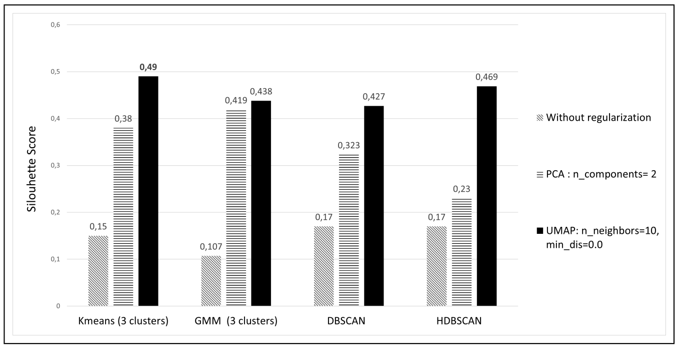

Fig. 3. Performance comparison of clustering techniques using all features based on Silhouette score for small dataset

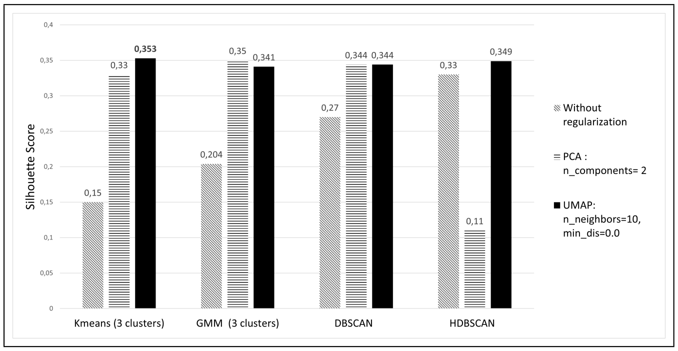

Fig. 4. Performance comparison of clustering techniques using all features based on Silhouette score for a large dataset

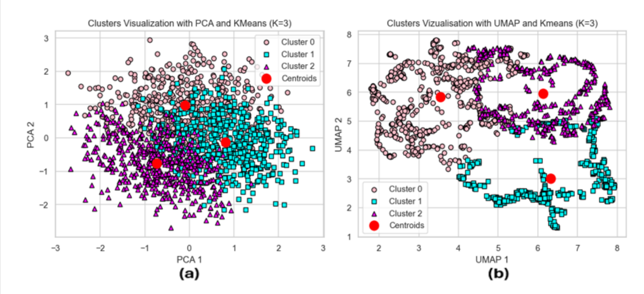

Fig. 5. (a), and (b) K-means partitions for k=3, respectively with PCA and UMAP regularization for large dataset

9500) and faster computation time (5.08s). In contrast, the baseline admitted only 85% of critical patients. For the large dataset, critical admission dropped to 20% due to capacity limitations but still doubled the baseline performance (10%). Additionally, the clustering-based approach lowered the cost from 26,500 (baseline) to 21,840, despite increased complexity and computation time (15.70s vs. 12.90s). These results highlight the value of integrating a clustering phase to group patients by severity, allowing more effective prioritization and better use of limited resources, especially under high-demand conditions.

TABLE II

COMPARISON OF OPTIMIZATION RESULTS WITH AND WITHOUT CLUSTERING

	Clustering results	Average of the Optimization Results		
	% Critical Patient (Cluster)	% Critical Admission	COST	CPU-Time (s)
Short set	30%-40%	100%	7220	5.08
Large set	30%-40%	20%	21840	15.70
Short set (Baseline)	–	85%	9500	4.20
Large set (Baseline)	–	10%	26500	12.90

VI. CONCLUSION

This paper proposes an innovative two-stage solution for the Hospital Patient Bed Assignment Problem (HPBAP) during emergencies, using unsupervised machine learning and optimization techniques. In the first stage, patients are clustered according to their contagious states into three groups, using a K-means algorithm with the UMAP regularization method. In the second stage, an optimization model is applied to minimize assignment costs by efficiently assigning patients to beds and managing patient transfers. The proposed approach demonstrates strong performance in prioritizing critical patients, particularly for smaller datasets. For larger datasets, future work will explore hybrid methods, combining metaheuristics, supervised and unsupervised techniques, to improve scalability and assignment outcomes.

REFERENCES

[1] Demeester, Peter, et al. *A hybrid tabu search algorithm for automatically assigning patients to beds.* Artificial Intelligence in Medicine 48.1 (2010): 61-70.

[2] Vancroonenburg, W., D. Goossens, and F. Spieksma. *On the complexity of the patient assignment problem.* Tech. rep., KAHO Sint-Lieven, Gebroeders De Smetstraat 1, Gent, Belgium (2011).

[3] Schäfer, F., Walther, M., Grimm, D. G., and Hübner, A. *Combining machine learning and optimization for the operational patient-bed assignment problem.* Health Care Management Science, (2023), 26(4), 785-806.

[4] Hong, W., Adrian D.H., and R. A. Taylor. *Predicting hospital admission at emergency department triage using machine learning.* PloS one 13.7 (2018): e0201016.

[5] L. Yan, H.-T. Zhang, and Y. Xiao. *Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan,* medRxiv, 2020.

[6] H. Yao, N. Zhang, and R. Zhang. *Severity detection for the coronavirus disease 2019 (COVID-19) patients using a machine learning model based on the blood and urine tests* Frontiers in Cell and Developmental Biology, (2020), vol. 8, pp. 1–10.

[7] I., Sadegh, A., Amir, et N. E. Akhondzadeh. *Cluster-based analysis of COVID-19 cases using self-organizing map neural network and K-means methods to improve medical decision-making.* Informatics in medicine unlocked, 2022, vol. 32, p. 101005.

[8] BOUTAZART, Y., SATORI, H., HAMIDI, M., et al. *Covid-19 dataset clustering based on K-means and EM algorithms.* International Journal of Advanced Computer Science and Applications, 2023, vol. 14, no 3.

[9] Benito-León, J., Del Castillo, M. D., Estirado, A., Ghosh, R., Dubey, S., and Serrano, J. I. *Using unsupervised machine learning to identify age- and sex-independent severity subgroups among patients with COVID-19: observational longitudinal study.* Journal of Medical Internet Research, 23(5), e25988.

[10] Y. LEBNI, J., ABBAS, J., MORADI, F., et al. *How the COVID-19 pandemic effected economic, social, political, and cultural factors: A lesson from Iran.* International Journal of Social Psychiatry, 2021, vol. 67, no 3, p. 298-300.

[11] Chaudhary L., Singh B., *Community detection using unsupervised machine learning techniques on COVID-19 dataset.* Social Network Analysis and Mining 2021;11(1):1–9.

[12] Health Organization *World. Coronavirus disease (COVID-19) pandemic. Geneva,* World Health Organization; 2020

[13] JEDIDI, H., BEN-ROMDHANE, H., NOUAOURI, I., and KRICHEN, S., *Bi-objective Hospital Bed Assignment Problem in Emergencies.* International COnference on Decision Aid and Artificial Intelligence (ICODAI 2024). Atlantis Press, 2025. p. 136-146.

[14] JEDIDI, H.,NOUAOURI, I., BEN-ROMDHANE, H., and KRICHEN, S., . *Min-Conflict Heuristic Approach for Elective Patient Bed Assignment Problem.* 2024 10th International Conference on Control, Decision and Information Technologies (CoDIT). IEEE, 2024. p. 1424-1429.

[15] JEDIDI, H., BEN-ROMDHANE, H., NOUAOURI, I., and KRICHEN, S., *A two-stage approach combining machine learning and optimization for the hospital patient bed assignment problem in emergencies.* Procedia Computer Science, 2024, vol. 246, p. 4316-4324.

[16] AL-MEKHLAFI, A., KLAWITTER, S., et KLAWONN, F.,*Standardization with zlog values improves exploratory data analysis and machine learning for laboratory data.* Journal of Laboratory Medicine, 2024, vol. 48, no 5, p. 215-222.

[17] Ester, M., Kriegel, H. P., Sander, J., and Xu, X. *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.* Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD),1996.