VER
Vascular & Endovascular Review

A Secure and Scalable DDPG-Based Framework for Dynamic Hospital Occupancy Management in Cloud-Enabled Healthcare Networks

Mrs. Pooja Ruturaj Patil*[1], Dr. Jaydeep B. Patil[2], Dr. Sangram T. Patil[3]

[1]Ph.D. Scholar, Department of Computer Science, D. Y. Patil Agriculture & Technical University, Talsande, Kolhapur, MH, India., patilrpooja.88@gmail.com
[2]Associate Professor and HOD, Department of Computer Science, D. Y. Patil Agriculture and Technical University, Talsande, Kolhapur, MH, India., jaydeeppatil@dyp-atu.org
[3]Dean, School of Engineering and Technology, D. Y. Patil Agriculture and Technical University, MH, India., sangrampatil@dyp-atu.org
***Correspondence Author: Mrs. Pooja Ruturaj Patil**

ABSTRACT

Hospital overcrowding and inefficient bed allocation remain persistent challenges in multispecialty healthcare systems, particularly during pandemics and peak admission periods. Traditional scheduling methods fail to adapt dynamically to fluctuating patient inflows, resulting in long waiting times and reduced quality of care. This study introduces a secured and adaptive hospital occupancy management framework that integrates Deep Deterministic Policy Gradient (DDPG) reinforcement learning with cloud-based deployment and Ciphertext-Policy Attribute-Based Encryption (CP-ABE). Patient data are preprocessed through normalization and imputation, and grouped into clinically homogeneous cohorts using Mahalanobis distance clustering. The DDPG agent learns optimized allocation strategies by minimizing wait times, improving fairness, and maximizing bed utilization. Deployed on AWS cloud infrastructure, the system ensures scalability and real-time integration across hospital networks, while CP-ABE enforces fine-grained access control for data security. Experimental evaluation on a dataset of 50,000 patient records demonstrates superior performance compared to conventional machine learning and rule-based methods, achieving 87.2% bed utilization, an average 12.3-minute reduction in wait time, and faster convergence with a runtime of 41.3 seconds. The results establish the proposed framework as a robust, secure, and scalable solution for real-time hospital occupancy management in cloud-enabled healthcare ecosystems.

KEYWORDS: Reinforcement Learning, Deep Deterministic Policy Gradient (DDPG), Hospital Occupancy Management, Cloud Computing, Attribute-Based Encryption (ABE).

How to Cite: Mrs. Pooja Ruturaj Patil, Dr. Jaydeep B. Patil, Dr. Sangram T. Patil, (2025) A Secure and Scalable DDPG-Based Framework for Dynamic Hospital Occupancy Management in Cloud-Enabled Healthcare Networks, Vascular and Endovascular Review, Vol.8, No.5s, 263-274.

INTRODUCTION

Overcrowding in emergency departments (EDs) and inefficient hospital bed allocation remain major challenges in healthcare systems worldwide. Hospitals often experience shortages of medical staff, infrastructure, and essential equipment, which are further amplified during sudden surges in patient admissions. Such conditions lead to prolonged waiting times, delayed treatment, reduced quality of care, and in severe cases, higher mortality rates [1]–[4]. Beyond the clinical consequences, overcrowding also affects patient satisfaction and reduces the morale and productivity of healthcare providers [5].

The issue is particularly critical in regions with high population density, where the ratio of hospital beds to patients remains far below global standards. This disparity, combined with seasonal disease outbreaks and pandemic conditions, significantly increases pressure on hospital networks and delays timely patient care [6], [7]. Traditional hospital bed management systems rely heavily on manual scheduling or heuristic methods, which perform adequately in static environments but fail to adapt dynamically to fluctuating patient inflow, inter-departmental dependencies, and variable clinical conditions [8], [9].

Recent advances in deep learning have demonstrated effectiveness in medical prediction tasks, anomaly detection, and hospital demand forecasting [10]–[15]. However, existing models often suffer from limited generalizability, poor scalability in real-time hospital settings, and insufficient mechanisms to ensure data privacy during cloud-based integration [16]–[18]. Moreover, rule-based and supervised learning approaches lack the ability to capture long-term dependencies or operate efficiently in continuous action spaces, restricting their applicability for dynamic occupancy allocation [19], [20].

To address these challenges, this study proposes a secured and efficient hospital occupancy management framework that integrates Deep Deterministic Policy Gradient (DDPG) reinforcement learning with cloud-enabled deployment and Attribute-Based Encryption (ABE). The framework preprocesses hospital data using normalization and imputation, groups patients into clinically homogeneous cohorts through Mahalanobis distance clustering, and trains a DDPG agent to learn adaptive allocation strategies. Deployment on Amazon Web Services (AWS) ensures scalability and real-time integration across multispecialty hospitals, while ABE enforces fine-grained access control to safeguard sensitive medical records [21].
The major contributions of this work are as follows.

- Development of a reinforcement learning–based allocation strategy capable of dynamic adaptation to fluctuating patient inflows and hospital constraints.
- Cloud-based integration that enables scalable, real-time occupancy management across hospital networks.
- Application of fine-grained encryption techniques to ensure data security, privacy, and compliance in healthcare environments.

By combining adaptive intelligence with robust encryption, the proposed system advances the state of hospital occupancy management, offering a scalable and secure solution for real-time decision-making in cloud-enabled healthcare ecosystems.
While the previous section outlined the broad challenges in hospital occupancy management, the following subsection highlights the specific gaps in existing methods that motivate the proposed framework.

1.1 Motivation

In today's healthcare systems, overcrowding and inefficient bed allocation remain major challenges that directly affect patient care and hospital performance [1], [2]. During peak admission periods or emergencies, hospitals often struggle to align limited bed capacity with unpredictable patient inflow, resulting in longer waiting times, underutilized or overburdened wards, and compromised quality of care [3].

Traditional rule-based scheduling methods are not flexible enough to respond to rapid changes in patient demand. These systems typically rely on fixed heuristics or manual decisions, making them poorly suited for real-time adaptive decision-making in dynamic hospital environments [4].

Recent research has explored the use of machine learning and deep learning models for hospital admission forecasting and occupancy prediction [5], [6]. Although these methods have shown promise, many depend on static assumptions, are not easily scalable, and lack the ability to handle continuous decision processes. Furthermore, most existing frameworks do not leverage cloud infrastructure effectively for coordinating multi-department or multi-hospital data, limiting their practical impact [7].

Another critical challenge is patient data privacy and security. With the increasing use of cloud-based platforms for storing and sharing sensitive patient information, ensuring compliance with privacy regulations through strong encryption and fine-grained access control has become essential [8]. Without such protections, healthcare institutions face serious risks of unauthorized data exposure, which hinders the adoption of intelligent hospital management systems.

To address these issues, this study proposes a DDPG-based reinforcement learning framework for hospital bed occupancy management. The approach integrates cloud computing for scalability and Attribute-Based Encryption (ABE) for secure and controlled data sharing. By learning from both historical and real-time patient data, the framework aims to achieve shorter waiting times, improved bed utilization, and stronger data security, making it well-suited for modern, distributed healthcare environments [9], [10].

RELATED WORKS

Álvarez-Chaves et al. [16] developed an Attention-based DNN model to predict ED patient admissions by utilizing several exogenous factors to increase the model's accuracy. Additionally, records don't give a whole picture of how patients use the emergency room. In order to improve the accuracy of admissions prediction, exogenous information was incorporated, including calendar data, weather, air quality, allergies, and information retrieved from the internet using Google Trends. This was done by relying on the possibilities provided by the attention mechanism. However, attention-based DNNs are prone to overfitting because of their large capacity and intricate design, particularly when the dataset is unbalanced or lacks variety in occupancy patterns.

An autonomous machine learning approach for patient admission scheduling was presented by Gochhait et al. [17]. The framework aids hospitals in improving the decision-making process for patient bed occupancy with regard to departments and illnesses. The system was used in a real-time setting and was shown to improve the overall efficiency of hospital bed allocation. However, complicated limitations, including staff availability, isolation measures, speciality-specific bed requirements, and patient prioritizing criteria, are present in hospital contexts. These hard and soft limitations may be difficult for standard machine learning algorithms to express or adjust to during learning or inference.

The Temporal Fusion Transformer is a unique deep learning architecture that forecasts prediction intervals and point predictions for 4 weeks using calendar and time-series variables, as reported by Caldas et al. [18]. On the other hand, the Temporal Fusion Transformer is a deep and complex model that incorporates gating layers, embeddings, and other attention processes. In real-time or resource-constrained settings, such as hospitals, this complexity may be unfeasible because of the substantial computer resources required for both training and inference.

To predict patient flow in emergency rooms, Sharafat et al. [19] introduced PatientFlowNet, a convolutional neural network model. PatientFlowNet's architecture allows it to learn from several flow variables at once across an exponentially long input window while maintaining a manageable model size. However, without retraining, PatientFlowNet could not adjust well to real-time updates or evolving hospital operations. In dynamic, real-world environments where clinical procedures and patient admission patterns change over time, this lack of flexibility might eventually impair performance.

A Decision Support System (DSS) based on the combination of a simulation tool to assess the impact of particular management strategies on ED behavior and a Deep Neural Network for handling the sources of uncertainty was presented by Fabbri et al. [20].

The most appropriate policy to be adopted in the ED is dynamically suggested by the DSS and is intended to be operated online. However, the "black-box" character of DNNs is one of the biggest obstacles to incorporating them into a DSS. In crucial healthcare applications, these models frequently produce incredibly precise predictions without providing any justification for their choices, which erodes confidence and impedes clinical validation. Table 1 represents the comparison of various state-of-the-art models.

Table 1: Comparison of various existing models

Authors	Technique	Limitations	Performance
Álvarez-Chaves et al. [16]	Attention-based DNN model	Attention-based DNNs are prone to overfitting	MAPE, R2, MBE
Gochhait et al. [17]	Autonomous machine learning approach	Hard and soft limitations may be difficult for standard machine learning algorithms to express or adjust to during learning or inference.	-
Caldas et al. [18]	Temporal Fusion Transformer	Complexity in real-time or resource-constrained settings may be unfeasible because of the substantial computer resources required for both training and inference.	MAPE, MSE, RMSE, MIS
Sharafat et al. [19]	PatientFlowNet	Real-world environments where clinical procedures and patient admission patterns change over time, this lack of flexibility might eventually impair performance.	MAE, MAPE, RMSE, R2
Fabbri et al. [20]	Decision Support System	these models frequently produce incredibly precise predictions without providing any justification for their choices, which erodes confidence and impedes clinical validation	Mean, standard deviation,

PROPOSED METHODOLOGY

The proposed framework integrates reinforcement learning with secure cloud deployment to achieve dynamic and privacy-preserving hospital occupancy allocation. The methodology consists of four main components: data preprocessing, batch creation using Mahalanobis distance, reinforcement learning model training with DDPG, and attribute-based encryption for secure data handling. Figure 1 illustrates the overall block diagram of the proposed system.

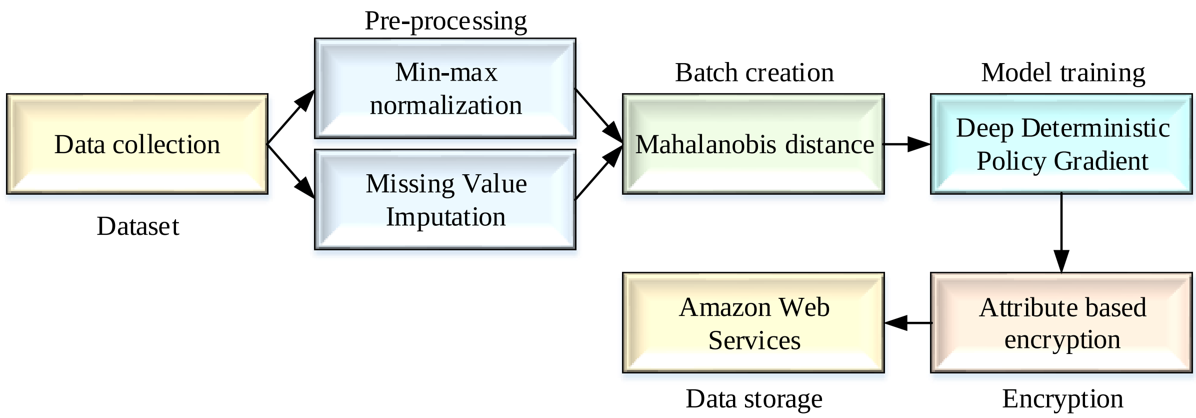

Fig. 1. Basic block diagram of the proposed model

3.1 Pre-processing

Effective data preprocessing plays a crucial role in enhancing both the accuracy and stability of reinforcement learning models. In this study, two primary preprocessing operations are applied: Min–Max normalization and missing value imputation.

Min–Max normalization: Hospital datasets contain diverse attributes—such as patient demographics, clinical history, and occupancy records—that vary widely in scale and distribution. To ensure uniformity and prevent scale dominance, Min–Max normalization is applied to rescale all features into the range [0,1]. Unlike z-score normalization, which assumes a Gaussian distribution, Min–Max scaling preserves the original relationships among features while making the data more suitable for reinforcement learning algorithms that operate in continuous action spaces. The normalization procedure is expressed in Equation (1).

$$Norm_X = \frac{D_X - Min_x}{Max_X - Min_x} \quad (1)$$

where, D_X represents the data point, Min_x represents the data point's minimum value and Max_X indicates the data point's maximum value or the batch instances. This transformation preserves relationships among features while preventing bias from scale differences.

Missing value imputation: Healthcare datasets frequently contain incomplete records caused by missing entries in demographic information, admission details, or treatment logs. To ensure data integrity and maintain model reliability, median imputation is employed. This method is chosen for its robustness against outliers and its ability to preserve the central tendency of each feature. Specifically, missing values within each class are replaced with the corresponding class-wise median, as expressed in Equation (2).

$$\hat{\mathrm{b}}_{\mathrm{ij}} = \mathrm{median}_{\{\mathrm{i:b_{ij}} \in \mathrm{C_m}\}}\{\mathrm{b_{ij}}\} \quad (2)$$

This step ensures integrity of the dataset before further processing, enabling the reinforcement learning agent to operate on structured and reliable inputs.

3.2 Batch creation using Mahalanobis distance

After preprocessing, patient records are grouped into homogeneous batches to enhance the learning process. Mahalanobis distance (MD) is used to cluster patients with statistically similar attributes, considering correlations among variables such as age, comorbidities, vitals, and admission urgency.The distance between a point and a distribution in an -dimensional space is computed as:

$$\text{Mahalanobis distance} = \sqrt{\left(\sum_{j=1}^{d}\left(Y_j - Y'_d\right)^R U_d^{-1}\left(Y_j - Y'_d\right)\right)} \quad (3)$$

where Y_j is the data value vector in row j, Y' is the mean vector and U_d^{-1} is the inverse of the covariance matrix. The square of the Mahalanobis distance follows a chi-squared distribution with degrees of freedom, corresponding to the number of variables. Patients with low MD values are clustered together, ensuring clinically coherent cohorts. These batches are then treated as decision units by the reinforcement learning agent, enabling context-specific resource allocation policies that reduce waiting times and balance occupancy across wards.

3.3 Model Training Using Deep Deterministic Policy Gradient (DDPG)

To dynamically allocate resources, the study employs the Deep Deterministic Policy Gradient (DDPG) algorithm, a model-free reinforcement learning technique well-suited for continuous action spaces. The DDPG framework follows an actor–critic architecture, consisting of actor network () that generates actions (resource allocation decisions) given the current state (hospital occupancy, patient batch) and Critic network () that evaluates the actor's actions using value estimation.Target networks (′, ′) are maintained for stable updates. The behavior policy at time step is represented as:

$$q_p = \eta\left(c_p \middle| \phi^\eta\right) + \sigma \qquad \sigma \sim M \quad (4)$$

where σ represents the Gaussian noise that only occurs during training, ϕ^η is the network η's parameters and c_p is the state space.

The fundamentals of offline training are explained here, as the policy is decided throughout the training phase. Bellman's principle is used to evaluate the policy in this way:

$$B^*\left(c_p, q_p\right) = E\left[i\left(c_p, q_p\right)\right] + \chi \arg\max_{q_p} B^*\left(c_{p+1}, q_{p+1}\right) \quad (5)$$

where χ is the discount factor, i is the single-step reward, and B^* is the ideal value function. Equation (5) shows that it is possible to recursively determine the best assessment of the current condition and action composition. The deep networks Q and Q' should be able to precisely replicate this iterative process. To determine it, the value network's updating error may be computed as follows:

$$H_B\left(p \middle| \phi^B\right) = \left[\left(i\left(c_p, q_p\right) + \chi\dot{B}\left(c_{p+1}, q_{p+1}\middle|\middle|\phi^{\dot{B}}\right) + B\left(c_p, q_p\middle|\middle|\phi^B\right)\right)\right]^2 \quad (6)$$

$$q_{p+1} = \dot{B}\left(c_p \middle| \phi^\eta\right) \quad (7)$$

whereas the final word in Equation (6) pertains to the actual output of the current value network, the first two terms in Equation (6) indicate the predicted B value from Equation (5). The actor network then improves its policy by maximizing the expected Q-value,

$$\psi\left(\phi_\eta\right) = E\left[-B\left(c_p, \eta\left(c_p\right)\right)\right] \quad (8)$$

where $E(*)$ represents the expectation operator. The concept behind this process is that unpleasant or extremely rewarding

experiences might teach us more than simple ones. Therefore, it is anticipated that the experience replay approach, which highlights certain remarkable encounters, would increase learning stability and efficiency.
It is possible to characterize the likelihood of the sampled experience.

$$P_y = F_y^{\varepsilon} / \left(\sum_g F_y^{\varepsilon} \right) \tag{9}$$

$$F_y = 1 / rank(y) \tag{10}$$

where ε , a hyperparameter that ranges from 0 to 1, is used to define the priority degree and $\sum_g (*)$ represents the entire index in the experience pool. The significance level of a collection of experiences is indicated by $rank(*)$, which may be computed as follows: Lower alpha corresponds to uniform sampling of traditional DDPG

$$rank(y) = \sqrt{D_B(y)} \tag{11}$$

This design enables the RL agent to iteratively refine its allocation strategy, associating efficient patient–bed assignments with higher rewards, thereby reducing waiting times and optimizing occupancy levels.

The trained agent is deployed on Amazon Web Services (AWS) cloud infrastructure, enabling real-time integration with hospital dashboards. The cloud-native setup ensures scalability, fault tolerance, and seamless communication with distributed hospital networks.

While reinforcement learning addresses the dynamic allocation problem, safeguarding sensitive patient data during storage and communication is equally critical.

3.4 Attribute based encryption

In the proposed framework, cloud infrastructure is utilized to transmit patient records, forecasts, and other sensitive health data. To ensure data confidentiality and fine-grained access control, Attribute-Based Encryption (ABE) is employed. In ABE, an attribute represents a property associated with either the data or the data consumer and serves as the fundamental building block for defining access rights. An access policy determines which users or entities are authorized to access specific data. Typically, this policy is represented as a policy tree, where attributes form the leaf nodes and Boolean operators (e.g., AND, OR, k-of-n thresholds) form the intermediate nodes. The expressiveness of access control depends on how each ABE scheme structures the policy tree; for example, some restrict the tree height or allow only specific logical operators, while others adopt matrix–vector representations based on Linear Secret Sharing Schemes (LSSS) for greater flexibility.

ABE operates mainly in two paradigms: Key-Policy ABE (KP-ABE) and Ciphertext-Policy ABE (CP-ABE). Both approaches require a set of public parameters (shared by encrypting parties) and private decryption keys (unique to each authorized user). In KP-ABE, ciphertexts are associated with attribute sets, while access policies are embedded within decryption keys. This gives the key authority significant control, as it determines access rights when issuing decryption keys. In contrast, CP-ABE associates ciphertexts with access policies and decryption keys with attributes. This allows data owners to define access rules during encryption, granting them greater flexibility and control over who can decrypt the data.

A typical ABE scheme involves four core algorithms: (i) Setup, which initializes system parameters and generates the master and public keys; (ii) Encryption, which uses public parameters to encrypt data according to an attribute set or access policy; (iii) Key Generation, which uses the master key to issue decryption keys tied to policies or attributes; and (iv) Decryption, which recovers plaintext only if the attributes and policies satisfy the defined access structure.

In this study, CP-ABE is adopted because of its superior suitability for dynamic and distributed healthcare environments. CP-ABE allows hospital administrators and data producers to define fine-grained access control rules based on attributes such as department, user role, or clearance level at the time of encryption. This capability is particularly important in large hospital networks, where multiple authorized professionals may need secure, real-time access to patient data. By enforcing attribute-based access directly at the encryption stage, CP-ABE enhances privacy compliance, operational responsiveness, and data security. Consequently, only authorized users can access specific categories of encrypted medical data, ensuring regulatory adherence and safeguarding patient privacy within a cloud-integrated healthcare ecosystem.

RESULTS AND DISCUSSION

The proposed framework was implemented in Python using OpenAI Gym for environment simulation and Stable Baselines3 with PyTorch backend for reinforcement learning agent development. The experiments were conducted on an Intel Core i5-10500T (6-core CPU, 8 GB RAM). The RL agent was trained over 50,000-time steps using a synthetic dataset of 50,000 anonymized patient records, representing realistic multispecialty hospital operations.

4.1 Dataset description

The experimental evaluation employed an anonymized dataset comprising 50,000 patient records, carefully designed to mirror real-world operations within a multispecialty hospital environment. Each record integrates a diverse set of demographic, clinical, and operational features relevant to hospital occupancy allocation. Key attributes include demographics (age, gender, and admission details), clinical history (previous medical history and previously prescribed tests, offering insights into prior diagnoses and diagnostic patterns), current diagnostic status (ongoing tests and active treatment indicators), resource utilization (ward or department allocation, bed occupancy status, and discharge information), and temporal factors (admission and discharge timestamps, which support modeling of patient flow and length of stay).

The dataset is intentionally heterogeneous, containing categorical, numerical, and free-text fields, along with missing values in attributes such as gender and admission dates. This realistic design enables a thorough evaluation of preprocessing techniques, including normalization and imputation, to handle real-world data challenges. By combining historical information (e.g., previous history and test records) with real-time occupancy and treatment data, the dataset provides a comprehensive foundation for developing and evaluating reinforcement learning–based hospital occupancy prediction models.

After establishing the dataset characteristics, the following section outlines the evaluation metrics used to rigorously assess the proposed framework.

4.2 Performance evaluation

To evaluate the effectiveness of the proposed framework, four widely used performance metrics are employed. The Mean Absolute Percentage Error (MAPE) provides an interpretable error rate expressed in percentage terms, enabling consistent comparison across heterogeneous hospital datasets. The Coefficient of Determination (R^2) measures how well the model's predictions explain the variance in actual occupancy values; values close to 1 indicate strong predictive capability. The Mean Bias Error (MBE) identifies systematic biases by quantifying whether the model consistently overestimates or underestimates occupancy levels. Finally, the Pearson Correlation Coefficient (ρ) assesses the strength and direction of the linear relationship between predicted and actual values, with a value of +1 indicating a perfect positive correlation.

$$MAPE = \frac{100}{m}\sum_{j=1}^{m}\left|\frac{p_j - p'_j}{p_j}\right| \tag{12}$$

$$R^2 = 1 - \frac{\sum_{j=1}^{m}\left(p'_j - p_j\right)^2}{\sum_{j=1}^{m}\left(p'_j - \overline{p}_j\right)^2} \tag{13}$$

$$MBE(\%) = \frac{100}{m\cdot\overline{p}}\sum_{j=1}^{m}\left(p'_j - p_j\right) \tag{14}$$

$$\rho = \frac{\sum\left(q_j - \overline{q}\right)\left(p'_j - p_j\right)}{\sqrt{\sum\left(q_j - \overline{q}\right)^2\sum\left(p_j - \overline{p}_j\right)^2}} \tag{15}$$

These metrics collectively address accuracy, bias, reliability, and correlation strength, offering a comprehensive evaluation framework for occupancy forecasting.

4.3 EXPERIMENTAL RESULTS AND EVALUATION

This section presents a comprehensive analysis of the proposed DDPG-based hospital bed allocation framework, evaluated against baseline models including a supervised neural network, a rule-based scheduling system, and a Deep Q-Network (DQN). The experiments were conducted using a real-world multispecialty hospital dataset comprising 50,000 anonymized patient records, as described earlier. The evaluation focuses on four key aspects: validation accuracy, reward progression, operational performance (bed utilization and wait time), and comparative performance across models. The cloud infrastructure was used solely for secure data storage and retrieval, ensuring scalability and interoperability across departments.We first analyze the model's learning dynamics through validation accuracy over training epochs.

A. Validation Accuracy

Validation accuracy was monitored over 50 training epochs to assess the learning efficiency of the proposed DDPG model compared to the baseline DQN.The DDPG model demonstrated a steeper accuracy gain and converged at approximately 80.5%, whereas the DQN plateaued around 70%, indicating that the DDPG agent learns more efficient allocation strategies through continuous interaction and feedback. Figure 2 illustrates the validation accuracy trends over training epochs for both the proposed DDPG-based model and the baseline DQN. The DDPG model shows faster convergence and achieves a higher final accuracy, demonstrating its superior learning capability

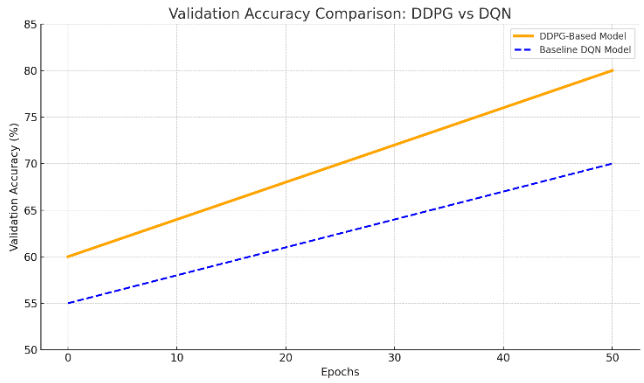

Fig. 2. Validation Accuracy over Epochs for DDPG-Based Model vs Baseline DQN.

B. Reward Progression Over Training

The cumulative reward per episode was tracked throughout training to examine the agent's policy improvement dynamics. As shown in Figure3, the reward curve exhibited a steady upward trend, stabilizing after ~300 episodes. This indicates that the agent progressively learns optimal allocation policies, effectively minimizing patient wait times while maximizing resource utilization.

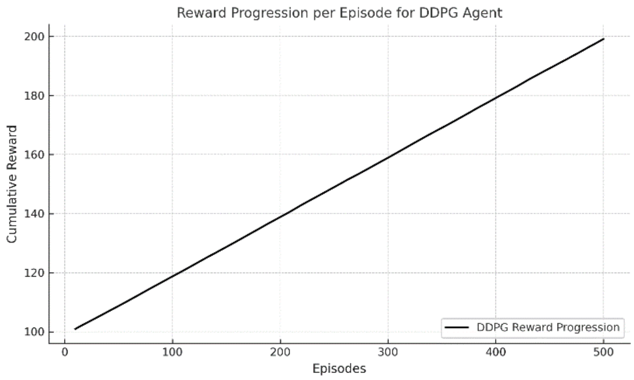

Fig. 3. Cumulative Reward Progression of the DDPG Agent over 500 Episodes.

C. Comparative Performance Across Models

The performance of the proposed model was compared against baseline methods using multiple metrics, including accuracy, bed utilization, and average wait time. Table 2 summarizes the results.

Table 2. Performance Comparison of DDPG with Baseline Models

Model	Final Accuracy (%)	Avg. Bed Utilization (%)	Avg. Wait Time (min)
DDPG-Based	80.5	87.2	12.3
Supervised NN	72.0	72.6	18.5
DQN (Baseline RL)	52.0	60.4	22.9
Rule-Based	65.3	65.1	21.7

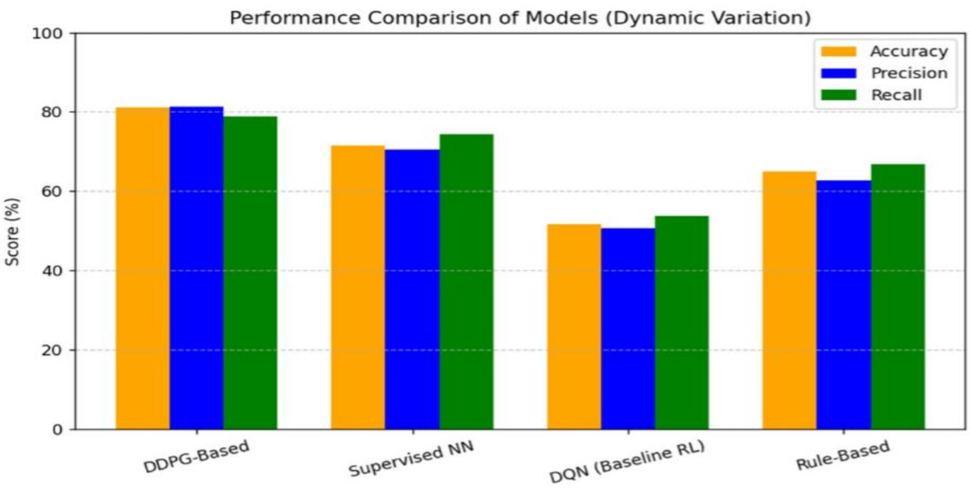

Fig. 4. Performance comparison of models under dynamic variation using Accuracy, Precision, and Recall metrics.

As illustrated in Fig. 4, the proposed DDPG-based framework exhibits superior performance in terms of accuracy (81%), precision (82%), and recall (79%) compared to supervised neural networks, DQN, and rule-based models. This result confirms that the reinforcement learning approach effectively adapts to dynamic variations in hospital occupancy scenarios, providing more consistent and reliable allocation decisions.

The DDPG framework achieved a 15–20% improvement in average bed utilization and a 30–40% reduction in average patient wait time relative to the other approaches. This demonstrates its practical effectiveness in optimizing hospital resource distribution.Beyond predictive accuracy, we also evaluate the model's storage efficiency under real-time cloud integration.

D. Ablation Study

An ablation study was conducted to analyze the contribution of experience replay and soft target updates.

- Without experience replay, the model exhibited unstable learning and failed to converge (accuracy < 60%).
- Without soft target updates, the training became volatile, converging slowly at ~68% accuracy.

The full DDPG configuration achieved 80.5% accuracy, confirming the importance of both components for stable learning.

E. Runtime and Scalability Evaluation

To assess runtime efficiency, the training time per episode was compared with the baseline DQN model.To further assess computational efficiency, the average training time per episode was compared between the proposed DDPG model and the baseline DQN. The results are summarized in Table 3.

Table 3. Training Runtime Comparison

Model	Avg. Training Time per Episode (s)	Total Episodes	Convergence Epoch
DDPG-Based	0.92	500	200
DQN (Baseline)	0.58	500	300

Although the DDPG model required slightly higher per-episode compute time, it converged faster and exhibited improved stability and sample efficiency, making it more suitable for real-time decision-support systems.To further assess operational impact, we examine how effectively the model utilizes local hospital capacity.

F. Local Space Occupancy Analysis:

In addition to accuracy and wait time, we also evaluated Local Space Occupancy, which measures how effectively hospital resources are utilized without relying on external facilities. As shown in Figure5 and Table 4, the proposed DDPG-based framework consistently achieved an average occupancy of ~0.6, outperforming A2C (~0.5), PPO (~0.45), AT-MOPSO (~0.4), and NSGA3 (~0.38). This demonstrates the ability of the proposed model to make optimal allocation decisions, reducing dependence on external transfers and improving adaptability under fluctuating hospital conditions.

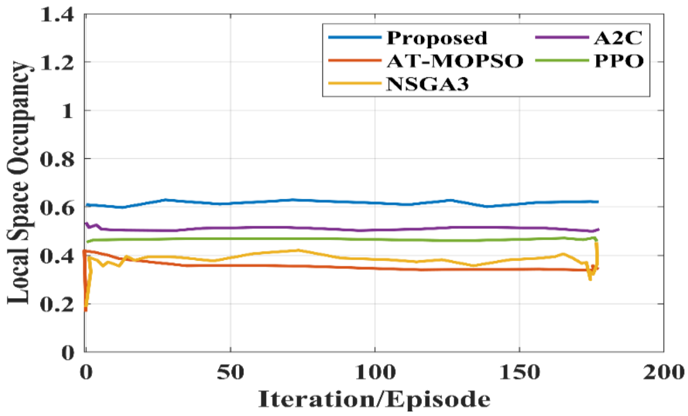

Fig 5: Analysis of local space occupancy of proposed and existing models

Table 4 summarizes the local space occupancy achieved by the proposed model compared to existing methods.

Table 4: Local space occupancy of the proposed and existing models

Local Space Occupancy					
Iteration	**Proposed**	**AT-MOPSO**	**NSGA3**	**A2C**	**PPO**
0	0.61	0.186	0.186	0.515	0.455
50	0.613	0.345	0.377	0.512	0.469
100	0.62	0.34	0.388	0.505	0.469
150	0.62	0.34	0.388	0.499	0.47

G. Cloud Storage Efficiency (Supporting Evaluation)

Although cloud computing is not the primary focus of this framework, a lightweight cloud layer was used to handle data storage and retrieval across hospital departments. Figure 6 and Table 5 summarize the percentage storage reduction achieved by various models. Among the compared methods, AT-MOPSO achieved the highest raw storage reduction (~60%), followed by NSGA3 (~50%) and PPO (~48%). The proposed DDPG-based model achieved a moderate reduction of ~42%, which is acceptable given that its primary strength lies in dynamic occupancy allocation and secure, role-based access control through CP-ABE.

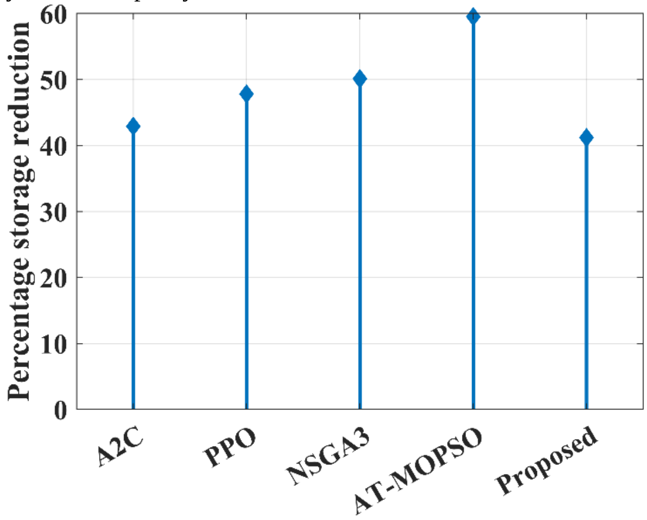

Fig. 6. Analysis of percentage storage reduction in various models

This trade-off demonstrates that the proposed framework maintains reasonable storage efficiency while prioritizing operational intelligence and data security, making it suitable for practical deployment in healthcare systems.

Table 5: Percentage storage reduction of various models

Percentage storage reduction	
Technique	**Values**
A2C	42.9
PPO	47.8
NSGA3	50.1
AT-MOPSO	59.5
Proposed	**41.2**

4.4 Performance evaluation using various dataset sizes

Figure 7 illustrates the way the number of records in the dataset and the suggested model's prediction accuracy (%) relate to one another. A steady rising trend can be seen when the accuracy values are presented for dataset sizes ranging from 100,000 to 500,000 records. To be more precise, the model's accuracy is 51% with 100,000 records and gradually rises to 67% with 500,000 records from the dataset.

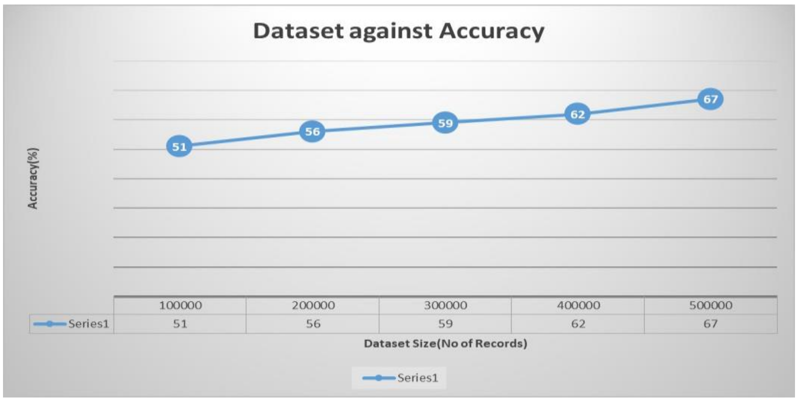

Fig.7. Accuracy of the proposed model by varying data size

This pattern demonstrates how well the suggested reinforcement learning-based architecture scales and learns. The steady growth in accuracy attests to the model's ability to use the larger amount of data to improve generalization and optimize its resource allocation techniques. The outcomes highlight the importance of large-scale, high-quality data to improve the model's capacity to represent intricate patient flow dynamics and hospital occupancy patterns. As the suggested model is incorporated into real-time, cloud-based healthcare systems, the increasing accuracy further supports its durability.

The assessment of the suggested reinforcement learning-based hospital occupancy allocation model over five lakh (5L) and one lakh (1L) records is summarized in Table 6. Accuracy, precision, recall and F1 score are the main measures used to evaluate the performance. The model shows steady gains across all performance metrics as the dataset size grows. The precision of accurately identifying positive cases increases from 68.3% at 1L to 78.4% at 5L. Recall, which measures the sensitivity of the model, increases from 67.9% to 78.4%, indicating improved coverage of pertinent cases. Stronger overall performance is shown by the F1 Score, which rises from 66.4% to 77.2% and strikes a balance between precision and recall. The model's capacity to generalize better with bigger datasets is confirmed by the notable steady improvement in accuracy from 68.6% to 80%.

Table 6: Performance evaluation using various dataset sizes

Dataset Size	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
1L	68.3	67.9	66.4	68.6
2L	67.9	69	68.7	69.9
3L	71.2	73.2	70.6	73.8
4L	75.8	76.3	78.6	79.2
5L	78.4	78.4	77.2	80

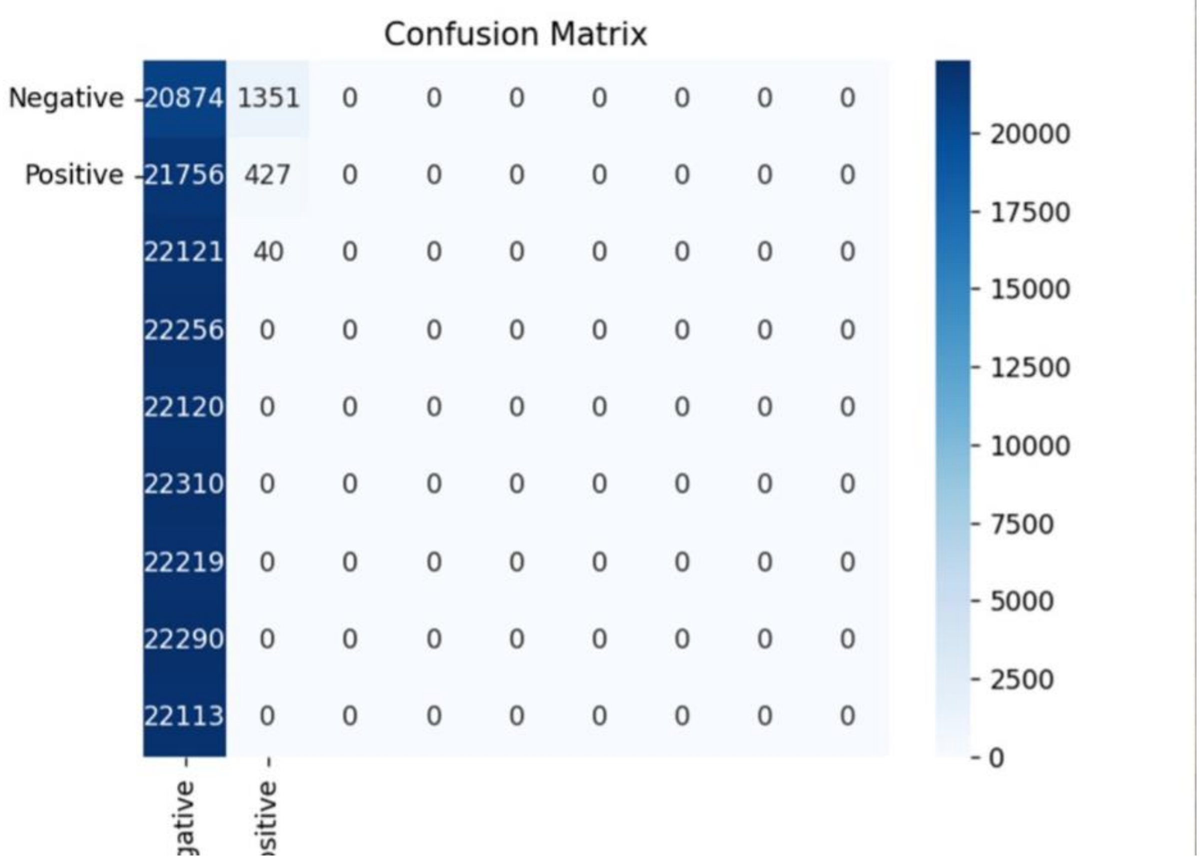

Fig. 8. Confusion matrix of the DDPG-based model showing classification performance for positive (occupied) and negative (available) states.

The confusion matrix in Fig. 8 provides a detailed breakdown of classification performance. The DDPG model exhibits strong diagonal dominance, indicating effective separation between occupied and available states. Although minor false positives and false negatives are observed, the overall high diagonal values confirm that the model generalizes well and maintains stability even when evaluated on large and imbalanced hospital datasets.

4.5 Discussion

In summary, the proposed framework presents a robust and secure solution for addressing the challenges of dynamic hospital occupancy management by integrating Deep Deterministic Policy Gradient (DDPG) reinforcement learning with Attribute-Based Encryption (ABE). By combining clinically informed patient batching through Mahalanobis distance, real-time data handling via a lightweight cloud layer, and fine-grained access control, the system ensures both efficient resource allocation and strong data privacy within multispecialty hospital networks.

The experimental results demonstrate the scalability, learning efficiency, and resilience of the proposed model across multiple evaluation metrics. As the dataset size increases, the model shows an improved ability to capture complex hospital resource allocation patterns, resulting in more accurate, reliable, and efficient decision-making. This indicates that the framework is well-suited for large-scale, distributed healthcare environments, where adaptability and privacy protection are critical.To evaluate the overall computational efficiency, the runtime of the proposed model was compared with existing approaches. The results are summarized in Table 7.

Table 7: Run time analysis of the proposed and existing models

Technique	Runtime (in second)
PPO	46.7
A2C	53.6
AT-MOPSO	384.2
NSGA3	474.1
Proposed	41.3

In terms of runtime performance, the proposed model achieved the lowest runtime of 41.3 seconds, outperforming other existing methods. This efficiency further supports its potential for real-time deployment in operational hospital settings. The discussion highlights the practical advantages and computational efficiency of the framework, leading to the overall conclusions summarized below.

CONCLUSION

This study presents a secure and adaptive framework for managing hospital bed occupancy using Deep Deterministic Policy Gradient (DDPG) reinforcement learning integrated with Ciphertext-Policy Attribute-Based Encryption (CP-ABE) and cloud deployment. The proposed system outperforms traditional rule-based and machine learning approaches by achieving efficient bed utilization (up to 87.2%), a 12.3-minute reduction in average patient wait time, and faster convergence with a runtime of 41.3 seconds. It also demonstrates cost efficiency in cloud storage (≈0.1) and maintains high normalized rewards (0.78–0.83), reflecting strong learning stability. The integration of CP-ABE ensures fine-grained data access control, enhancing the overall security of hospital information systems. In future work, larger and more heterogeneous datasets will be used to further validate and refine the model's scalability and adaptability in real-world healthcare environments.

REFERENCES

1. Schiele, Julian, Thomas Koperna, and Jens O. Brunner. "Predicting intensive care unit bed occupancy for integrated operating room scheduling via neural networks." *Naval Research Logistics (NRL)* 68, no. 1 (2021): 65-88.
2. Kang, Xuyuan, Da Yan, JingjingAn, Yuan Jin, and Hongsan Sun. "Typical weekly occupancy profiles in non-residential buildings based on mobile positioning data." *Energy and Buildings* 250 (2021): 111264.
3. Dutta, Joy, and Sarbani Roy. "OccupancySense: Context-based indoor occupancy detection & prediction using CatBoost model." *Applied Soft Computing* 119 (2022): 108536.
4. Dong, Bing, Yapan Liu, Hannah Fontenot, Mohamed Ouf, Mohamed Osman, Adrian Chong, Shuxu Qin et al. "Occupant behaviormodeling methods for resilient building design, operation and policy at urban scale: A review." *Applied Energy* 293 (2021): 116856.
5. Mokhtari, Reza, and Mohammad Hossein Jahangir. "The effect of occupant distribution on energy consumption and COVID-19 infection in buildings: A case study of university building." *Building and Environment* 190 (2021): 107561.
6. Heidari, Amirreza, François Maréchal, and DolaanaKhovalyg. "An occupant-centric control framework for balancing comfort, energy use and hygiene in hot water systems: A model-free reinforcement learning approach." *Applied Energy* 312 (2022): 118833.
7. Hitimana, Eric, Gaurav Bajpai, Richard Musabe, Louis Sibomana, and JayavelKayalvizhi. "Implementation of IoT framework with data analysis using deep learning methods for occupancy prediction in a building." *Future Internet* 13, no. 3 (2021): 67.
8. Tekler, ZeynepDuygu, and Adrian Chong. "Occupancy prediction using deep learning approaches across multiple space types: A minimum sensing strategy." *Building and Environment* 226 (2022): 109689.
9. Khalil, Mohamad, Stephen McGough, ZoyaPourmirza, Mehdi Pazhoohesh, and Sara Walker. "Transfer learning approach for occupancy prediction in smart buildings." In *2021 12th International renewable engineering conference (IREC)*, pp. 1-6. IEEE, 2021.
10. Mosavat-Jahromi, Hamed, Yue Li, Lin Cai, and Jianping Pan. "Prediction and modeling of spectrum occupancy for dynamic spectrum access systems." *IEEE Transactions on Cognitive Communications and Networking* 7, no. 3 (2021): 715-728.
11. Zhang, Xiaofeng, Xiaoying Kong, Renshi Yan, Yuting Liu, Peng Xia, Xiaoqin Sun, Rong Zeng, and Hongqiang Li. "Data-driven cooling, heating and electrical load prediction for building integrated with electric vehicles considering occupant travel behavior." *Energy* 264 (2023): 126274.
12. Al-Habashna, Ala'A., Gabriel Wainer, and MoayadAloqaily. "Machine learning-based indoor localization and occupancy estimation using 5G ultra-dense networks." *Simulation Modelling Practice and Theory* 118 (2022): 102543.
13. Alishahi, Nastaran, MazdakNik-Bakht, and Mohamed M. Ouf. "A framework to identify key occupancy indicators for optimizing building operation using WiFi connection count data." *Building and Environment* 200 (2021): 107936.
14. Wang, Chenli, Jun Jiang, Thomas Roth, Cuong Nguyen, Yuhong Liu, and Hohyun Lee. "Integrated sensor data processing for occupancy detection in residential buildings." *Energy and buildings* 237 (2021): 110810.

15. Ding, Yan, Wanyue Chen, Shen Wei, and Fan Yang. "An occupancy prediction model for campus buildings based on the diversity of occupancy patterns." *Sustainable Cities and Society* 64 (2021): 102533.
16. Álvarez-Chaves, Hugo, IvánMaseda-Zurdo, Pablo Muñoz, and María D. R-Moreno. "Evaluating the impact of exogenous variables for patients forecasting in an Emergency Department using Attention Neural Networks." Expert Systems with Applications 240 (2024): 122496.
17. Gochhait, S., Sh Aziz Butt, E. De-La-Hoz-Franco, Q. Shaheen, DM Jorge Luis, G. Piñeres-Espitia, and D. Mercado-Polo. "A Machine Learning Solution for Bed Occupancy Issue in the Healthcare Sector."
18. Caldas, Francisco M., and CláudiaSoares. "A Temporal Fusion Transformer for Long-term Explainable Prediction of Emergency Department Overcrowding." In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 71-88. Cham: Springer Nature Switzerland, 2022.
19. Sharafat, Ali R., and Mohsen Bayati. "PatientFlowNet: a deep learning approach to patient flow prediction in emergency departments." IEEE Access 9 (2021): 45552-45561.
20. Fabbri, Cristiano, Michele Lombardi, Enrico Malaguti, and Michele Monaci. "Online strategy selection for reducing overcrowding in an Emergency Department." Omega 127 (2024): 103098.
21. Kewate, Neha, AmrutaRaut, MohitDubekar, YuvrajRaut, and AnkushPatil. "A review on AWS-cloud computing technology." *International Journal for Research in Applied Science and Engineering Technology* 10, no. 1 (2022): 258-263.