Investigating the Impact of Ventilator Bundle Compliance Rates on Predicting ICU Patients with Risk for Hospital-Acquired Ventilator-Associated Pneumonia Infection in Saudi Arabia

Ghaida S. Alsaab[a] and Sarah A. Alkhodair[b]

Information Technology Department, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

Keywords: Hospital-Acquired Pneumonia, Machine Learning, Ventilator-Associated Pneumonia, Ventilator Bundle, ICU.

Abstract: Pneumonia is the most common infectious disease picked up in the Intensive Care Unit (ICU) and accounts for nearly 27% of all hospital infections—from 5% to 40% of ICU patients on mechanical ventilation risk getting infected by ventilator-associated pneumonia. Fortunately, by identifying patients more likely to contract pneumonia, up to 50% of ventilator-associated pneumonia infections can be avoided. To our knowledge, this is the first study that tackles the problem of identifying ICU patients with a high risk of developing ventilator-associated pneumonia in Saudi hospitals, considering the impact of ventilator bundle compliance rates on the predicted results. Five machine learning models were built using two real life datasets from the Health Electronic Surveillance Network (HESN) at the Saudi Ministry of Health. Results show that including ventilator bundle compliance rates data in the prediction process yields improved results in general; however, the extent of enhancement varies across models.

1 INTRODUCTION

Infections are a significant cause of concern in healthcare and can lead to severe illnesses and even death (Clinic, 2023a). Pneumonia, caused by a lung bacterial infection, is one of the most common infections, accounting for nearly 27% of all hospital infections (Clinic, 2023b). It is also a frequent occurrence in Intensive Care Units (ICUs) and is the leading cause of death from infectious diseases, killing millions of people every year (Coalition, 2022). Patients undergoing mechanical ventilation are particularly susceptible to a dangerous complication known as ventilator-associated pneumonia (VAP), with anywhere from 5 to 40% of these patients developing lung infections (Humayun et al., 2021). Fortunately, up to 50% of VAP infections can be prevented through the identification of patients who are at high risk of developing pneumonia and through the implementation of preventive bundles (Umscheid et al., 2011). The Ventilator Bundle, a set of interventions followed by ICUs in hospitals, is one example of such a preventative measure that has emerged as a pivotal factor in analyzing infection sources.

Several studies have investigated using machine learning (ML) techniques for predicting which patients get infected by various infectious diseases such as pneumonia by using historical clinical data, lab results, or X-rays. For instance, Yahyaoui and Yumuşak focused on predicting pneumonia and asthma using deep neural network (DNN) and K-nearest neighbors (KNN) (Yahyaoui and Yumuşak, 2021). Sun et al. developed two ML models: classification and regression trees (CART) and logistic regression (LR) using electronic health records to predict community-acquired pneumonia after respiratory tract infection (RTI) consultations in primary care (Sun et al., 2022). Chen et al. developed six ML models for predicting postoperative pneumonia in patients after liver transplantation: logistic regression (LR), support vector machine (SVM), random forest (RF), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), and gradient boosting machine (GBM). They reported the best performance, 73% accuracy and 61.8% sensitivity, was achieved by the XGBoost model (Chen et al., 2021). Abujaber et al. developed a decision tree (DT) model to predict ventilator-associated pneumonia (VAP) in patients with moderate to severe traumatic brain injury (TBI), achieving 83.5% accuracy,

[a] https://orcid.org/0009-0008-2378-5492
[b] https://orcid.org/0000-0001-8428-3092

Alsaab, G. and Alkhodair, S.
Investigating the Impact of Ventilator Bundle Compliance Rates on Predicting ICU Patients with Risk for Hospital-Acquired Ventilator-Associated Pneumonia Infection in Saudi Arabia.
DOI: 10.5220/0012574900003657

In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2024) - Volume 2*, pages 797-802
ISBN: 978-989-758-688-0; ISSN: 2184-4305

71% precision, 43% sensitivity, and a 54% F1-score (Abujaber et al., 2021). While these studies have employed machine learning to predict the onset of infectious diseases, none of them, to our knowledge, have investigated the potential impact of compliance with the ventilator bundle for developing pneumonia using machine learning techniques. Addressing this research gap presents a significant opportunity to gain better insights and understanding of the risk factors for pneumonia and other infectious diseases in the ICU, which can significantly enhance patient outcomes and reduce the incidence of infectious diseases, thus improving the quality of care for ICU patients.

To the best of our knowledge, this study is the first attempt to tackle the problem of investigating the potential impact of ventilator bundle compliance rates at Saudi hospitals for predicting ICU patients with a high risk for hospital-acquired ventilator-associated pneumonia (VAP). The main objective of this study is to build several machine learning (ML) models with and without consideration of ventilator bundle compliance rates and to compare their performance in predicting ICU patients at risk for VAP infection. We extensively tested five ML models on two real life datasets from the Saudi Ministry of Health's Electronic Surveillance Network (HESN). We reported our results using several prediction performance evaluation measures: accuracy, sensitivity, precision, and F1-score. The results obtained show that including ventilator bundle compliance rates data in the prediction process yields improved results in general; however, the extent of enhancement varies across models.

2 METHODOLOGY

Figure 1 illustrates the methodology followed in this study. The following subsections briefly cover each step.

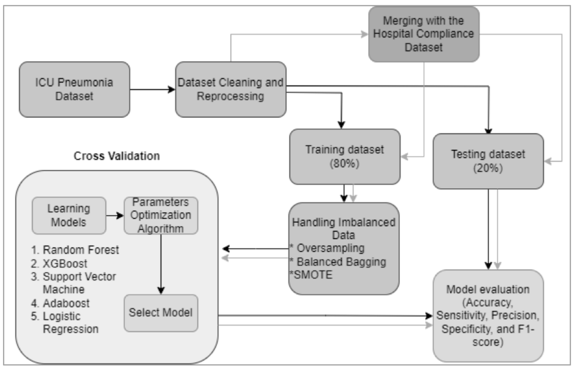

Figure 1: Methodology followed in this study.

2.1 Dataset Description

Two real-life datasets from the Saudi Ministry of Health's Electronic Surveillance Network (HESN) were collected and utilized for this study. The gathering process started by conducting several meetings and focus group sessions comprised of ML experts and healthcare professionals who provided practical insights, guiding the utilization of electronic medical records and ICU ventilator bundle data for the proposed model. Critical features for the predictive model were identified, including age, gender, ICU stay duration, underlying medical conditions, and ventilator bundle compliance. The collection process resulted in two meticulously gathered datasets for this study: the "ICU Pneumonia Dataset" and the "Hospitals Compliance Dataset."

2.1.1 The ICU Pneumonia Dataset

The ICU pneumonia dataset was designed to facilitate an in-depth investigation of pneumonia cases within the intensive care unit (ICU) environment by providing a comprehensive perspective on the patient population and their healthcare experiences within the ICU. This dataset covers data from 2017 through 2020.

2.1.2 The Hospitals Compliance Dataset

The Hospitals Compliance dataset describes the compliance of hospitals with the ventilator bundle. This dataset covers data from 2016 through 2022.

2.2 Data Cleaning and Preprocessing

Several steps of data preprocessing and preparation were conducted to ensure the coherence and relevance of the datasets to the current study's primary objective, which is investigating the potential impact of ventilator bundle compliance rates on predicting ventilator-associated pneumonia (VAP) among ICU patients; we only retained records of patients who used ventilators in the ICU. All records with missing target values were removed. Features with irregular cardinalities or exhibiting null values only were removed, such as "Central Line" and "BSI." Irrelevant features such as "Death Date," "BSIEventDate," "Total reviews," and "Overall bundle" were also removed. Furthermore, features providing duplicate information such as "Head Elevation," "Daily Sedation Hold," "PUD Prophylaxis," "DVT Prophylaxis," and "Daily Oral Care" were removed. Finally, all records from the years 2016, 2021, and 2022 were removed from the Hospitals Compliance dataset. These data cleaning and preprocessing steps resulted in 802,356 records and 12 features in the ICU pneumonia dataset as described in Table 1, and 12,250 records and 11 features in the Hospitals Compliance dataset as described in Table 2.

2.3 Merging the Two Datasets

Investigating the potential impact of ventilator bundle compliance rates on predicting ventilator-associated pneumonia (VAP) among ICU patients mandates comparing the prediction results of ML models trained only on the essential information in the "ICU Pneumonia dataset" as well as ML models trained on a dataset consisting of both the "ICU Pneumonia dataset" and the associated records from the "Hospitals Compliance dataset" that complements the primary dataset by sharing key features regarding the compliance to the ventilator bundle. The inner merging technique was employed, focusing on the shared key features such as region, hospital, unit, year, and month among the two datasets. This process resulted in a merged dataset comprising 77,577 records and 18 features.

2.4 Machine Learning Models

Five ML models were built and evaluated in our study: random forest (RF), support vector machine (SVM), logistic regression (LR), extreme gradient boosting machine (XGBoost), and adaptive boosting (AdaBoost). These models have been proven to achieve superior performance for similar research problems in the literature, especially binary classification problems with severe class imbalance (Khushi et al., 2021).

2.5 Experimental Settings

To investigate the impact of the hospital compliance with the ventilator bundle data on predicting VAP among ICU patients, we built and evaluated two models of each of the selected ML algorithms: one on the primary "ICU Pneumonia dataset" and one on the merged dataset. This resulted in the development of ten ML models for our experiment.

To build and evaluate the models, both datasets were split into an 80% training set and a 20% testing set. The Synthetic Minority Oversampling Technique (SMOTE) (Blagus and Lusa, 2013) was applied to the training set to handle the class imbalance issue before building the ML models. A fivefold cross-validation technique was employed to build the ML models, where each model experienced rigorous training and evaluation within the defined cross-validation framework, allowing for a comprehensive comparison of their respective performances, optimizing model selection, and enhancing their robustness and generalizability. Finally, the testing set was used to evaluate the prediction performance of the models. We reported our results using several performance evaluation measures: accuracy, sensitivity, precision, and F1-score.

The experiment was conducted using a HP computer with a Windows 64-bit operating system, a 2 GHz processor, and 8 GB of RAM. Tools used include Anaconda Navigator, Jupyter Notebooks, and pandas, NumPy, OS, and sklearn Python libraries.

3 RESULTS

Table 3 shows the obtained results of evaluating the prediction performance of the selected ML algorithms on the primary "ICU Pneumonia dataset" and the merged dataset in terms of accuracy, precision, sensitivity, specificity, and F1-score. Evaluation results showed that SVM emerged as the top-performing model in terms of accuracy, obtaining 89.48% accuracy using the primary dataset. Additionally, SVM demonstrated the highest sensitivity at 90.73%, showcasing its proficiency in correctly identifying positive instances. Precision, a critical metric for assessing the correctness of positive predictions, was notably high for both RF and XGBoost, reaching 97% on the merged dataset. Moreover, when considering the F1-score, XGBoost outperformed other models with a score of 92% on the merged dataset. It is noteworthy that random forest (RF) and XGBoost consistently excelled in recall, achieving the highest value of 89% on the merged dataset. Logistic regression (LR) demonstrated its strength in achieving specificity, with 74.62% on the merged dataset, signifying its proficiency in correctly identifying negative instances.

4 DISCUSSION

SVM, LR, RF, XGBoost, and AdaBoost are popular ML models proven to achieve high performance in many domains, including healthcare. We built and evaluated these five ML models on two datasets.

We found that LR yielded the worst overall prediction performance compared to the other models. More specifically, LR performed the worst in accuracy and sensitivity when the merged dataset is used, while it performed the worst in precision and F1-score with the primary dataset. We noticed that using the

Table 1: The primary dataset: The ICU Pneumonia dataset.

Feature	Explanation
Region	The region where hospital located
Hosp.	Hospital Name
Unit	ICU ID
Year	Specifies the year of the collected data
Month	Specifies the month of the collected data
Age	Patient age
Gender	Patient gender
Stay	Length of stay in the ICU in days
Central Line	The number of days the patient used Central Line to get medicines, blood, or nutrition
BSI	The number of days the patient got infected by Bloodstream Infection
Ventilator	The number of days the patient used Ventilator
Pneumonia	The number of days the patient got infected by Pneumonia

Table 2: The second dataset: The hospitals compliance to the ventilator bundle dataset.

Feature	Explanation
Region	The region where hospital located
Hosp.	Hospital Name
Unit	ICU ID
Year	Specifies the year of the collected data
Month Number	Specifies the month of the collected data
Head Elevation Rate	Calculated as (Head Elevation / total reviews) *100
Daily Sedation Hold Rate	Calculated as (Daily Sedation Hold / total reviews) *100
PUD Prophylaxis Rate	Calculated as (PUD Prophylaxis / total reviews) *100
DVT Prophylaxis Rate	Calculated as (DVT Prophylaxis / total reviews) *100
Daily Oral Care Rate	Calculated as (Daily Oral Care / total reviews) *100
Overall Compliance rate	Calculated as (overall bundle / total reviews) * 100

merged data with LR slightly decreases its performance by 1.4% in accuracy and 0.2% in sensitivity, while significantly improving its performance by 23% in precision, 26% in F1-score, and 6.1% in specificity. The significant improvement in some evaluation measures compared with the slight decrease in others with the merged dataset suggests that including the compliance with ventilator bundle data did improve the overall prediction performance of LR.

SVM, on the other hand, yielded the best overall prediction performance compared to the other models. However, it performed best in accuracy, sensitivity, and F1-score when using the primary dataset while achieving a comparable precision (96%) compared to the best performance achieved by RF (97%). We also noticed that using the merged data with SVM decreases its performance by 4.3% in accuracy, 4.6% in sensitivity, and 3% in F1-score, suggesting that the inclusion of the compliance with ventilator bundle data did not improve the prediction results of SVM.

We also noticed that the effect of including the compliance with ventilator bundle data on the prediction performance of AdaBoost was similar to that of LR. Using the merged data with AdaBoost did not affect its precision, slightly decreased its performance by 0.8% in sensitivity and 3% in specificity, while significantly improved its performance by 33% in precision and 34% in F1-score, suggesting that including the compliance with ventilator bundle data significantly improved the overall prediction performance of AdaBoost.

RF showed no significant improvements in its prediction performance when using the merged dataset compared to the primary dataset. The performance of RF slightly increased by 2.5% in accuracy, 2.9% in sensitivity, and 2% in F1-score, suggesting a slight impact of the ventilator bundle compliance data on its prediction performance. Similarly, XGBoost reflects a minimal decrease of 0.3% in accuracy and sensitivity and 0.9% in specificity with the merged dataset, suggesting a slight impact of including the compliance with ventilator bundle data on its prediction performance.

It is justifiable to argue that the merged dataset yields improved results; however, the extent of enhancement varies across models. AdaBoost and LR,

Table 3: The evaluation results of the prediction performance of the ML models on the two datasets.

Model	Accuracy		Precision		Sensitivity		F1-score		Specificity	
	Primary Dataset	Merged Dataset	Primary Dataset	Merged Dataset	Primary Dataset	Merged Dataset	Primary Dataset	Merged Dataset	Primary Dataset	Merged Dataset
LR	77.3%	75.9%	53%	76%	77.4%	77.2%	50%	76%	68.5%	74.6%
AdaBoost	83.1%	85.7%	53%	86%	83.7%	82.9%	52%	86%	58.4%	55.4%
RF	80.7%	83.2%	97%	97%	81%	83.9%	87%	89%	65.7%	54.7%
SVM	89.5%	85.2%	96%	96%	90.7%	86.1%	93%	90%	35.7%	45.5%
XGBoost	89.2%	88.9%	97%	97%	90.1%	89.8%	92%	92%	49.2%	48.3%

in particular, showcase significantly heightened overall prediction performance. AdaBoost stands out as the most profitable model when using the merged dataset. It significantly improved its precision and F1-score, highlighting the effectiveness of using the compliance rate with ventilator bundle data to achieve better predictions. LR comes next, where considering the ventilator bundle compliance rates also significantly improved its precision and F1-score. Both models balance precision and sensitivity by achieving a high F1-score, which is crucial for applications where both aspects are fundamental. RF also shows slightly heightened overall prediction performance using the merged dataset, while XGBoost yielded the least affected performance with the merged dataset compared to the primary dataset. Nevertheless, considering how differently the selected ML models responded to the merged dataset, additional experiments are required for a deeper and more comprehensive investigation of the potential impact of ventilator bundle compliance rates on predicting ventilator-associated pneumonia (VAP) among ICU patients in Saudi hospitals.

This study demonstrates the importance of the ventilator bundle compliance rates in monitoring and evaluating the risk of VAP among ICU patients. It examines the importance of ML in investigating the impact of ventilator bundle compliance rates at hospitals in improving the prediction of ICU patients at risk of VAP. It paves the pathway for further investigation and systematic application of machine learning and deep learning for improving the ventilator bundle compliance rates in ICU settings.

5 CONCLUSIONS

Pneumonia continues to pose a significant risk to patients in healthcare settings, particularly in the ICU. A significant proportion of these infections can be prevented by applying preventive bundles on identified patients with a risk of developing ventilator-associated pneumonia (VAP). Machine learning models have shown promise in identifying such patients; however, existing literature does not investigate the impact of ventilator bundle compliance rates on the prediction performance of such models. The current study addressed this gap by investigating the impact of ventilator bundle compliance rates on the performance of five ML models in predicting high-risk ICU patients for VAP in Saudi hospitals. Two real-life datasets were used to build and evaluate the models using several performance evaluation measures. The results highlight the potential impact of the ventilator bundle compliance rates on improving the prediction of ventilator-associated pneumonia (VAP) among ICU patients in Saudi hospitals. Nevertheless, additional experiments are required, considering that the extent of enhancement in the prediction performance varies across models.

ACKNOWLEDGEMENTS

We thank the Ministry of Health (MOH) in Saudi Arabia for providing the necessary data to conduct this research. We also thank Ms. Hessah Alasker, the Director of Advanced Analytics for Public Health at the Data and Artificial Intelligence Department at MOH, and Dr. Sattam bin Rushud, an ICU physician.

REFERENCES

Abujaber, A., Fadlalla, A., Gammoh, D., Al-Thani, H., and El-Menyar, A. (2021). Machine learning model to predict ventilator associated pneumonia in patients with traumatic brain injury: the c. 5 decision tree approach. *Brain Injury*, 35(9):1095–1102.

Blagus, R. and Lusa, L. (2013). SMOTE for high-dimensional class-imbalanced data. *BMC bioinformatics*, 14:106.

Chen, C., Yang, D., Gao, S., Zhang, Y., Chen, L., Wang, B., Mo, Z., Yang, Y., Hei, Z., and Zhou, S. (2021). Development and performance assessment of novel machine learning models to predict pneumonia after liver transplantation. *Respiratory Research*, 22(1):1–12.

Clinic, M. (2023a). Infectious diseases - symptoms & causes. https://www.mayoclinic.org/diseases-conditions/infectious-diseases/symptoms-causes/syc-20351173. (Accessed 13-10-2023).

Clinic, M. (2023b). Pneumonia - Symptoms and causes. https://www.mayoclinic.org/diseases-conditions/pneumonia/symptoms-causes/syc-20354204. (Accessed 13-10-2023).

Coalition, E. B. C. (2022). Stop Pneumonia / Every Breath Counts. https://stoppneumonia.org/issues/. (Accessed on 01/13/2023).

Humayun, T., Alshanbari, N., Alanazi, A., Aldecoa, Y. S., Alanazi, K. H., Saleh Bin, G., et al. (2021). Rates of ventilator-associated pneumonia in Saudi Ministry of Health hospitals; a two-year multi-center study. *Am J Infect Dis Micro*, 9:25–31.

Khushi, M., Shaukat Dar, K., Mahboob Alam, T., Hameed, I., Uddin, S., Luo, S., Yang, X., and Reyes, M. (2021). A comparative performance analysis of data resampling methods on imbalance medical data. *IEEE Access*.

Sun, X., Douiri, A., and Gulliford, M. (2022). Applying machine learning algorithms to electronic health records to predict pneumonia after respiratory tract infection. *Journal of Clinical Epidemiology*, 145:154–163.

Umscheid, C. A., Mitchell, M. D., Doshi, J. A., Agarwal, R., Williams, K., and Brennan, P. J. (2011). Estimating the proportion of healthcare-associated infections that are reasonably preventable and the related mortality and costs. *Infection Control & Hospital Epidemiology*, 32(2):101–114.

Yahyaoui, A. and Yumuşak, N. (2021). Deep and machine learning towards pneumonia and asthma detection. In *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, pages 494–497. IEEE.