

УПОТРЕБА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ ЗА ПРЕДИКЦИЈУ ИСХОДА
ТЕШКИХ БАКТЕРИЈСКИХ ИНФЕКЦИЈА НА ОДЕЉЕЊИМА ИНТЕНЗИВНЕ НЕГЕ

Марко Кими Милић, Шћепан Синановић

Висока медицинска школа струковних студија „Милутин Миланковић”, Београд, Србија

USE OF ARTIFICIAL INTELLIGENCE FOR PREDICTING OUTCOMES
OF SEVERE BACTERIAL INFECTIONS IN INTENSIVE CARE UNITS

Marko Kimi Milić, Šćepan Sinanović

High Medical College of Professional Studies “Milutin Milankovic”, Belgrade, Serbia

Сажетак

Циљ ове студије је да се процени ефикасност модела вештачке интелигенције (ВИ) у предвиђању исхода тешких бактеријских инфекција код пацијената на одељењима интензивне неге (ОИН). Главни фокус је на поређењу предвиђања заснованих на ВИ са традиционалним системима скоринга, укључујући APACHE II и SOFA, у погледу морталитета, дужине хоспитализације (LOS) и терапијског одговора. Спроведена је ретроспективна кохортна анализа на основу података пацијената са ОИН примљених у периоду од јануара 2020. до децембра 2024. године. Предобрада података укључивала је обраду недостајућих вредности методом MICE, кодирање категоријских варијабли и нормализацију континуираних варијабли. Развијена су четири модела машинског учења: XGBoost, Random Forest, дубоке неуронске мреже (DNN) и логистичка регресија. Перформансе модела процењиване су помоћу AUROC за предвиђање морталитета, MAE за предвиђање дужине боравка и F1-скора за процену терапијског одговора. XGBoost је показао највећу тачност у предвиђању морталитета, са AUROC вредношћу од 0,93, а затим следе Random Forest (0,91) и DNN (0,89). Логистичка регресија је имала најслабије перформансе (AUROC 0,82). Старост пацијената, SOFA скор и ниво лактата идентификовани су као најзначајнији предиктори. У предвиђању дужине боравка, XGBoost и Random Forest поново су показали боље резултате од осталих модела, са MAE вредностима од 3,2 и 3,4 дана. Када је у питању терапијски одговор, XGBoost је остварио највиши F1-скор од 0,81. ВИ модели су доследно надмашивали традиционалне системе скоринга. Модели вештачке интелигенције, посебно XGBoost, значајно побољшавају тачност предвиђања морталитета, дужине хоспитализације и терапијског одговора у поређењу са конвенционалним методама. Интеграција ових модела у клиничку праксу ОИН може унапредити доношење одлука и управљање пацијентима, омогућавајући персонализоване и правовремене интервенције. Препоручујемо додатну валидацију у различитим клиничким окружењима.

Кључне речи: вештачка интелигенција, одељења интензивне неге, машинско учење, бактеријске инфекција

Abstract

This study aims to evaluate the performance of artificial intelligence (AI) models in predicting outcomes of severe bacterial infections among intensive care unit (ICU) patients. The primary focus is to compare AI-based predictions with traditional scoring systems, including APACHE II and SOFA, in terms of mortality, length of stay (LOS), and therapeutic response. A retrospective cohort analysis was conducted using data from ICU patients admitted between January 2020 and December 2024. Data preprocessing involved handling missing values using the MICE method, encoding categorical variables, and normalizing continuous variables. Four machine learning models were developed: XGBoost, Random Forest, Deep Neural Networks (DNN), and Logistic Regression. Model performance was evaluated using AUROC for mortality prediction, MAE for LOS prediction, and F1-score for therapeutic response prediction. XGBoost demonstrated the highest accuracy in predicting mortality, with an AUROC of 0.93, followed by Random Forest (0.91) and DNN (0.89). Logistic Regression showed the lowest performance (AUROC 0.82). Age, SOFA score, and lactate level were identified as the most significant predictors. In predicting LOS, XGBoost and Random Forest again outperformed other models, with MAE values of 3.2 and 3.4 days, respectively. For therapeutic response, XGBoost achieved the highest F1-score of 0.81. AI models consistently outperformed traditional scoring systems. AI models, particularly XGBoost, significantly improve the accuracy of predicting mortality, LOS, and therapeutic response compared to conventional methods. Integrating these models into ICU practice could enhance clinical decision-making and patient management, promoting more personalized and timely interventions. Further validation in diverse clinical settings is recommended.

Keywords: artificial intelligence, intensive care units, machine learning, bacterial infections, prognosis

Увод

Тешке бактеријске инфекције остају водећи узрок морбидитета и морталитета међу пацијентима на одељењима интензивне неге (ОИН) широм света. Упркос напретку у антимикробним терапијама и супортивној нези, предви-

Introduction

Severe bacterial infections remain a leading cause of morbidity and mortality among patients in intensive care units (ICUs) worldwide. Despite advancements in antimicrobial therapies and supportive care, predicting outcomes

ђање исхода као што су морталитет, дужина боравка у болници и одговор на лечење остаје изазовно због сложености и варијабилности клиничких презентација [1]. У том контексту, вештачка интелигенција (ВИ) се појавила као обећавајући алат за побољшање клиничког доношења одлука кроз анализу великих и сложених скупова података [2].

Недавне студије су показале потенцијал вештачке интелигенције у раном откривању и прогнозирању сепсе, тешког и по живот опасног одговора на инфекцију. На пример, модел машинског учења који су развили Гао и др. постигао је површину испод криве оперативних карактеристика пријемника (AUROC) од 0,94 у предвиђању морталитета од сепсе, надмашујући традиционалне системе бодовања. Овај модел је користио смањени скуп карактеристика, побољшавајући интерпретабилност и клиничку применљивост [1].

Штавише, модели вештачке интелигенције су коришћени за предвиђање прикладности емпиријског лечења антибиотицима код пацијената на интензивној нези са болничким инфекцијама. Голдшмит и др. су развили модел машинског учења који је предвидео неприкладно емпиријско лечење антибиотицима са AUROC-ом од 0,773, омогућавајући ранију интервенцију и потенцијално побољшање исхода пацијената [2].

Интеграција вештачке интелигенције у клиничке токове рада нуди потенцијал за револуцију у лечењу пацијената омогућавањем ранијих интервенција и персонализованих стратегија лечења. Међутим, хетерогеност популација на интензивној нези и варијабилност етиологија инфекција представљају значајне изазове за развој и имплементацију модела. Идентификовање најрелевантнијих клиничких и лабораторијских предиктора, као и одабир робусних и интерпретабилних модела вештачке интелигенције, кључно је за клиничку применљивост [2].

Хипотеза

Претпостављамо да примена напредних модела вештачке интелигенције може значајно побољшати тачност предвиђања исхода код пацијената са тешким бактеријским инфекцијама на одељењима интензивне неге у поређењу са традиционалним алатима за процену ризика.

Примарни циљ ове студије је развој и валидација модела предвиђања заснованих на вештачкој интелигенцији који могу прецизно предвидети морталитет, дужину боравка и терапијски одговор код пацијената на интензивној нези који пате од тешких бактеријских инфекција.

such as mortality, length of hospital stay, and response to treatment remains challenging due to the complexity and variability of clinical presentations [1]. In this context, artificial intelligence (AI) has emerged as a promising tool to enhance clinical decision-making through the analysis of large and complex data sets [2].

Recent studies have demonstrated the potential of AI in early detection and prognosis of sepsis, a severe and life-threatening response to infection. For instance, a machine learning model developed by Gao et al. achieved an area under the receiver operating characteristic curve (AUROC) of 0.94 in predicting sepsis mortality, outperforming traditional scoring systems. This model utilized a reduced set of features, enhancing interpretability and clinical applicability [1].

Furthermore, AI models have been employed to predict the appropriateness of empiric antibiotic treatments in ICU patients with hospital-acquired infections. Goldschmidt et al. developed a machine learning model that predicted inappropriate empiric antibiotic treatment with an AUROC of 0.773, enabling earlier intervention and potentially improving patient outcomes [2].

The integration of AI into clinical workflows offers the potential to revolutionize patient management by enabling earlier interventions and personalized treatment strategies. However, the heterogeneity of ICU populations and the variability of infection aetiologies pose significant challenges to model development and implementation. Identifying the most relevant clinical and laboratory predictors, as well as selecting robust and interpretable AI models, is crucial for clinical utility [2].

Hypothesis

We hypothesize that the application of advanced AI models can significantly improve the accuracy of predicting outcomes in patients with severe bacterial infections in ICUs compared to traditional risk assessment tools.

The primary aim of this study is to develop and validate AI-based prediction models that can accurately forecast mortality, length of stay, and therapeutic response among ICU patients suffering from severe bacterial infections.

Research Questions

This study explores which artificial intelligence algorithms are most effective in predicting outcomes of severe bacterial infections in intensive care units, identifies the clinical and laboratory variables that most significantly contribute

Истраживачка питања

Ова студија истражује који су алгоритми вештачке интелигенције најефикаснији у предвиђању исхода тешких бактеријских инфекција на одељењима интензивне неге, идентификује клиничке и лабораторијске варијабле које најзначајније доприносе предиктивној тачности ових модела и упоређује перформансе AI модела са конвенционалним прогностичким системима бодовања који се обично користе у интензивној нези.

Ова студија има за циљ да попуни тренутну празнину у знању развојем робусног оквира заснованог на вештачкој интелигенцији, прилагођеног предвиђању исхода тешких бактеријских инфекција у условима интензивне неге. Резултати би могли олакшати рану стратификацију ризика и омогућити усмереније клиничке интервенције, што би на крају побољшало прогнозу пацијената.

Методе

Ова студија је осмишљена као ретроспективна кохортна анализа усмерена на развој и валидацију модела вештачке интелигенције (ВИ) за предвиђање исхода тешких бактеријских инфекција код пацијената на одељењима интензивне неге (ОИН). С обзиром на природу студије, која подразумева анализу потпуно анонимизованих, рутински прикупљених електронских здравствених картона (ЕЗК) без директне интеракције са пацијентом или прикупљања нових података, етичко одобрење није потребно. Овај приступ је у складу са важећим прописима и смерницама које се односе на секундарну употребу здравствених података у истраживачке сврхе.

Дизајн студије и прикупљање података

Студија користи рутински прикупљене податке од пацијената на интензивној нези примљених између јануара 2020. и децембра 2024. године из више медицинских центара како би се осигурала разноликост и генерализација. Процес прикупљања података биће строго ограничен на анонимизоване картоне пацијената како би се одржала приватност и поштовали прописи о заштити података. Критеријуми за укључивање ће обухватати одрасле пацијенте (старије од 18 година) са потврђеном дијагнозом тешке бактеријске инфекције, као што су сепса, бактеријемја или тешка бактеријска упала плућа, који су примљени на интензивну негу најмање 48 сати. Критеријуми за искључивање ће обухватати пацијенте са примарним вирусним или гљивичним инфекцијама без бактеријског учешћа, оне са значајним празнинама у подацима и поновне пријеме у истом

to the predictive accuracy of these models, and compares the performance of AI models with conventional prognostic scoring systems commonly used in critical care.

This study aims to fill the current knowledge gap by developing a robust AI-based framework tailored to the prediction of severe bacterial infection outcomes in ICU settings. The findings could facilitate early risk stratification and enable more targeted clinical interventions, ultimately improving patient prognosis.

Methods

This study is designed as a retrospective cohort analysis aimed at developing and validating artificial intelligence (AI) models to predict the outcomes of severe bacterial infections among patients in intensive care units (ICUs). Given the nature of the study, which involves the analysis of fully anonymized, routinely collected electronic health records (EHRs) without direct patient interaction or the collection of new data, ethical approval is not required. This approach aligns with current regulations and guidelines related to the secondary use of healthcare data for research purposes.

Study Design and Data Collection

The study utilizes routinely collected data from ICU patients admitted between January 2020 and December 2024 from multiple medical centres to ensure diversity and generalizability. The data collection process will be strictly confined to anonymized patient records to maintain privacy and comply with data protection regulations. Inclusion criteria will consist of adult patients (aged 18 years and older) with a confirmed diagnosis of severe bacterial infection, such as sepsis, bacteraemia, or severe bacterial pneumonia, who were admitted to the ICU for at least 48 hours. Exclusion criteria will include patients with primary viral or fungal infections without bacterial involvement, those with significant data gaps, and readmissions within the same hospitalization period.

The primary outcomes to be predicted are:

1. Mortality: Defined as in-hospital death during the ICU stay.
2. Length of Stay (LOS): The duration from ICU admission to either discharge or death.
3. Therapeutic Response: Improvement or deterioration within the first seven days of ICU admission, based on objective changes in vital signs, laboratory parameters (such as reduction in CRP, normalization of lactate), and overall clinical assessment.

The collected variables will include demographic charac-

периоду хоспитализације.

Примарни исходи које треба предвидети су:

1. Морталитет: дефинише се као смрт у болници током боравка на интензивној нези.
2. Дужина боравка (LOS): трајање од пријема на интензивну негу до отпуста или смрти.
3. Терапеутски одговор: побољшање или погоршање у првих седам дана од пријема на интензивну негу, на основу објективних промена виталних знакова, лабораторијских параметара (као што су смањење CRP-а, нормализација лактата) и опште клиничке процене.

Прикупљене варијабле ће обухватати демографске карактеристике (старост, пол, коморбидитете као што су дијабетес и хипертензија), клиничке податке (виталне знаке при пријему, SOFA резултате, резултате на Глазговој скали коме (GCS), лабораторијске резултате (број белих крвних зрнаца, CRP, прокалцитонин, лактат, маркере бубрежне функције), микробиолошке налазе (идентификоване патогене и профиле резистенције на антибиотике) и терапијске интервенције (врста и трајање антибиотика, употреба механичке вентилације и вазопресорска подршка). Ове варијабле су изабране на основу њихове клиничке релевантности и утврђене прогностичке вредности у окружењу интензивне неге [1, 2].

Претходна обрада података

Претходна обрада података ће укључивати управљање недостајућим вредностима коришћењем методе вишеструке импутације ланчаним једначинама (MICE) како би се одржао интегритет података и минимизирала пристрасност. Категоричке променљиве ће бити кодирани коришћењем једнократног кодирања, а континуиране променљиве ће бити нормализоване како би се осигурала униформност у целом скупу података. Поред тога, избор карактеристика ће се вршити коришћењем методе најмањег апсолутног скупљања и оператора селекције (LASSO) како би се смањила димензионалност уз очување најинформативнијих предиктора. Овај корак је кључан за побољшање тачности модела и смањење ризика од прекомерног уклапања, посебно имајући у виду сложену и вишедимензионалну природу података прикупљених из јединице интензивне неге.

Развој и евалуација модела

Неколико алгоритама машинског учења биће коришћено за развој предиктивних модела. Случајна шума и екстремно градијентно појачавање (XGBoost) биће коришћени због њихове робусности и способности да

teristics (age, gender, comorbidities such as diabetes and hypertension), clinical data (vital signs at admission, SOFA scores, Glasgow Coma Scale (GCS) scores), laboratory results (white blood cell count, CRP, procalcitonin, lactate, renal function markers), microbiological findings (identified pathogens and antibiotic resistance profiles), and therapeutic interventions (type and duration of antibiotics, use of mechanical ventilation, and vasopressor support). These variables were chosen based on their clinical relevance and established prognostic value in ICU settings [1, 2].

Data Preprocessing

Data preprocessing will involve the management of missing values using the Multiple Imputation by Chained Equations (MICE) method to maintain data integrity and minimize bias. Categorical variables will be encoded using one-hot encoding, and continuous variables will be normalized to ensure uniformity across the dataset. Additionally, feature selection will be performed using the Least Absolute Shrinkage and Selection Operator (LASSO) method to reduce dimensionality while preserving the most informative predictors. This step is crucial to enhance model accuracy and mitigate the risk of overfitting, particularly given the complex and multidimensional nature of ICU data.

Model Development and Evaluation

Several machine learning algorithms will be employed to develop predictive models. Random Forest and Extreme Gradient Boosting (XGBoost) will be utilized for their robustness and ability to handle non-linear interactions. Deep Neural Networks (DNNs) will also be explored to capture complex relationships, while Logistic Regression will serve as a baseline due to its interpretability and prevalence in clinical research.

Model training will be conducted using a 70-30 train-test split, with stratified sampling to ensure balanced representation of mortality and non-mortality cases. Cross-validation (5-fold) will be applied to validate model stability and reduce overfitting. Model performance will be assessed using the following metrics:

- Mortality Prediction: Area Under the Receiver Operating Characteristic Curve (AUROC)
- Length of Stay Prediction: Mean Absolute Error (MAE)
- Therapeutic Response Prediction: F1-score and precision-recall curve analysis

To ensure the robustness of the developed models, sensitivity analyses will be performed by varying the data preprocessing techniques and testing different feature selec-

обrade нелинеарне интеракције. Дубоке неуронске мреже (DNN) ће такође бити истражене за хватање сложених односа, док ће логистичка регресија служити као основа због своје интерпретабилности и распрострањености у клиничким истраживањима.

Обука модела ће се спроводити коришћењем поделе тренинг-тест 70–30, са стратификованим узорковањем како би се осигурала уравнотежена заступљеност случајева морталитета и неморталитета. Унакрсна валидација (петорострука) ће се применити како би се потврдила стабилност модела и смањило прекомерно прилагођавање. Перформансе модела ће се процењивати коришћењем следећих метрика:

- Предвиђање морталитета: површина испод криве радних карактеристика пријемника (AUROC)
- Предвиђање дужине боравка: средња апсолутна грешка (MAE)
- Предвиђање терапијског одговора: анализа F1-скора и криве прецизности и подсећања.

Да би се осигурала робусност развијених модела, анализе осетљивости биће спроведене варирањем техника претходне обраде података и тестирањем различитих метода избора карактеристика. Поред тога, анализе подгрупа ће истражити потенцијалне варијације у перформансама модела на основу старости пацијента, оптерећења коморбидитетом и врсте инфекције.

Софтвер и статистичка анализа

Сва претходна обрада података, развој модела и статистичке анализе биће изведени коришћењем Пајтона (верзија 3.x), користећи разне специјализоване библиотеке. *Pandas* ће се користити за манипулацију и чишћење података, док ће *NumPy* обрађивати нумеричке операције. За обуку модела машинског учења биће коришћени *Scikit-learn* и *XGBoost*. Модели дубоких неуронских мрежа биће развијени коришћењем *TensorFlow/Keras*-а због њихове флексибилности и скалабилности. Визуелизација резултата, укључујући ROC и криве прецизности и присећања, биће генерисана коришћењем *Matplotlib*-а и *Seaborn*-а.

Статистичке анализе, укључујући тестове за дистрибуцију података и анализу корелације, биће спроведене коришћењем *SciPy* и *Statsmodels*. Да би се осигурала репродуктивност, сви скриптови ће бити одржавани под контролом верзија коришћењем *Git*-а, а прорачуни ће се извршавати на кластеру високоперформансних рачунара како би се управљало великом количином података са јединица интензивне неге. Токови рада анализе биће документовани коришћењем *Jupyter Notebooks*-а

tion methods. Additionally, subgroup analyses will explore potential variations in model performance based on patient age, comorbidity burden, and infection type.

Software and Statistical Analysis

All data preprocessing, model development, and statistical analyses will be performed using Python (version 3.x), leveraging a variety of specialized libraries. *Pandas* will be used for data manipulation and cleaning, while *NumPy* will handle numerical operations. For machine learning model training, *Scikit-learn* and *XGBoost* will be employed. Deep neural network models will be developed using *TensorFlow/Keras* due to their flexibility and scalability. Visualization of results, including ROC and precision-recall curves, will be generated using *Matplotlib* and *Seaborn*.

Statistical analyses, including tests for data distribution and correlation analysis, will be conducted using *SciPy* and *Statsmodels*. To ensure reproducibility, all scripts will be maintained under version control using *Git*, and computations will be executed on a high-performance computing cluster to manage the large volume of ICU data. Analysis workflows will be documented using *Jupyter Notebooks* for transparency and reproducibility.

By utilizing a robust software framework, the study ensures that the analysis is transparent, reproducible, and adheres to the highest standards of computational research. The integration of machine learning techniques within a structured analytical pipeline aims to produce reliable and clinically relevant predictions.

Results

In this study, we developed and evaluated multiple AI models to predict ICU mortality, length of stay (LOS), and therapeutic response in patients with severe bacterial infections. The models tested included Random Forest (RF), Extreme Gradient Boosting (XGBoost), Deep Neural Networks (DNN), and Logistic Regression (LR) as a baseline. The primary metric for mortality prediction was the Area Under the Receiver Operating Characteristic Curve (AUROC), while LOS was evaluated using Mean Absolute Error (MAE). Therapeutic response was assessed using F1-score and precision-recall curves, given the potential class imbalance.

Patient Characteristics

A total of 2,000 patients with severe bacterial infections admitted to ICUs between January 2020 and December 2024 were included in the analysis. The mean age of the

ради транспарентности и репродуктивности.

Коришћењем робусног софтверског оквира, студија осигурава да је анализа транспарентна, репродукцибилна и да се придржава највиших стандарда рачунарског истраживања. Интеграција техника машинског учења у структурирани аналитички процес има за циљ да произведе поуздана и клинички релевантна предвиђања.

Резултати

У овој студији, развили смо и проценили више модела вештачке интелигенције за предвиђање морталитета, дужине боравка (LOS) и терапијског одговора на интензивној нези код пацијената са тешким бактеријским инфекцијама. Тестирани модели су укључивали случајну шуму (RF), екстремно градијентно појачавање (*XGBoost*), дубоке неуронске мреже (DNN) и логистичку регресију (LR) као основу. Примарна метрика за предвиђање морталитета била је површина испод криве оперативних карактеристика пријемника (AUROC), док је LOS процењен коришћењем средње апсолутне грешке (MAE). Терапеутски одговор је процењен коришћењем F1-скора и кривих прецизности и подсећања, с обзиром на потенцијални дисбаланс у класама.

Карактеристике пацијената

У анализу је укључено укупно 2000 пацијената са тешким бактеријским инфекцијама, примљених на одељења интензивне неге између јануара 2020. и децембра 2024. године. Просечна старост кохорте била је $62,4 \pm 14,7$ година, са претежно мушким полом (60%). Најчешћи коморбидитети били су хипертензија (45%), дијабетес мелитус (30%) и хронична болест бубрега (20%). Респираторне инфекције биле су најчешће место примарне инфекције (40%), затим инфекције крвотока (30%) и инфекције абдомена (20%).

Поређење преживелих и непреживелих открило је значајне разлике у демографским и клиничким варијаблама. Непреживели су били старији (просечна старост 67,8 година наспрам 60,2 године за преживеле) и имали су већу преваленцију коморбидитета, посебно хроничне болести бубрега (28% наспрам 18%). Поред тога, инфекције крвотока биле су чешће међу непреживелима (35% наспрам 28%), што указује на потенцијалну тежину повезану са бактеријемјом. Ови налази су у складу са претходним студијама које су наглашавале повећани ризик код старијих пацијената са коморбидитетима [1, 2].

Демографске и клиничке карактеристике пацијената на интензивној нези приказане су у табели 1. Ова табела

cohort was 62.4 ± 14.7 years, with a male predominance (60%). The most frequent comorbidities were hypertension (45%), diabetes mellitus (30%), and chronic kidney disease (20%). Respiratory infections were the most common primary infection site (40%), followed by bloodstream infections (30%) and abdominal infections (20%).

Comparing survivors and non-survivors revealed significant differences in demographic and clinical variables. Non-survivors were older (mean age 67.8 years vs. 60.2 years for survivors) and had a higher prevalence of comorbid conditions, particularly chronic kidney disease (28% vs. 18%). Additionally, bloodstream infections were more common among non-survivors (35% vs. 28%), indicating the potential severity associated with bacteraemia. These findings are consistent with previous studies emphasizing the increased risk among elderly patients with comorbidities [1, 2].

The demographic and clinical characteristics of ICU patients are presented in Table 1. This table highlights the significant differences between survivors and non-survivors, particularly in age and the prevalence of chronic conditions, which are key factors influencing ICU outcomes.

истиче значајне разлике између преживелих и непреживелих, посебно у старости и преваленцији хроничних стања, што су кључни фактори који утичу на исход на интензивној нези.

Табела 1. Демографске и клиничке карактеристике пацијената на интензивној нези

Карактеристика <i>Characteristic</i>	Укупно (n=2000) <i>Total (n=2000)</i>	Преживели (n=1500) <i>Survivors (n=1500)</i>	Непреживели (n=500) <i>Non-Survivors (n=500)</i>	р-вредност <i>p-value</i>
Старост (године) <i>Age (years)</i>	62,4 ± 14,7	60,2 ± 13,9	67,8 ± 15,3	<0,001
Мушки (%) <i>Male (%)</i>	60	58	65	0,03
Хипертензија (%) <i>Hypertension (%)</i>	45	40	55	<0,01
Дијабетес мелитус (%) <i>Diabetes Mellitus (%)</i>	30	28	37	0,02
Хронична болест бубрега (%) <i>Chronic Kidney Disease (%)</i>	20	18	28	<0,01
Респираторна инфекција (%) <i>Respiratory Infection (%)</i>	40	42	35	0,08
Инфекција крвотока (%) <i>Bloodstream Infection (%)</i>	30	28	35	0,05
Абдоминална инфекција (%) <i>Abdominal Infection (%)</i>	20	18	25	0,03
Мешовито/нејасно порекло (%) <i>Mixed/Unclear Origin (%)</i>	10	12	5	0,04

Table 1. Demographic and Clinical Characteristics of ICU Patients

Перформансе модела

Перформансе развијених модела вештачке интелигенције у предвиђању mortalитета на интензивној нези сумиране су у табели 2. XGBoost је постигао највиши AUROC (0,93), а одмах затим га је пратио Random Forest (0,91). DNN модел је постигао умерено добре резултате са AUROC-ом од 0,89, док је Логистичка регресија имала најниже резултате са AUROC-ом од 0,82. Ови резултати указују на то да су ансамбл методе (XGBoost и Random Forest) посебно ефикасне за предвиђање mortalитета због њихове способности да моделирају сложене обрасце података.

Табела 2. AUROC различитих модела за предвиђање mortalитета

Модел / Model	AUROC
XGBoost	0,93
Случајна шума / Random Forest	0,91
Дубока неуронска мрежа / Deep Neural Network	0,89
Логистичка регресија / Logistic Regression	0,82

Графикон 1 илуструје упоредне перформансе четири модела, јасно показујући супериорност XGBoost-а.

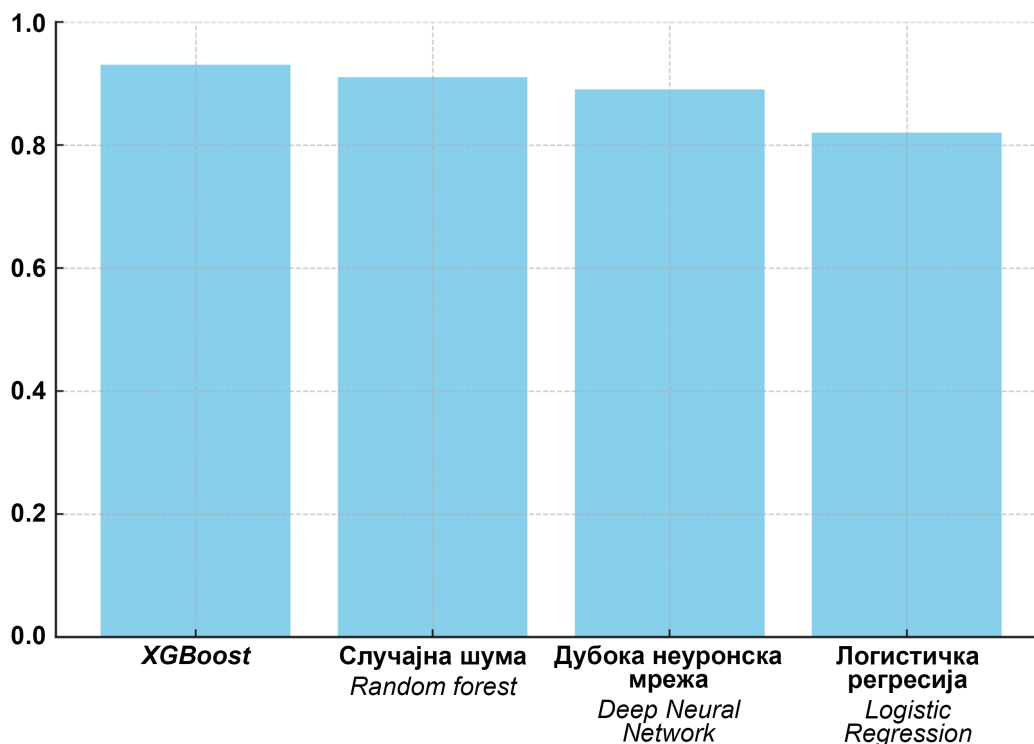
Model Performance

The performance of the developed AI models in predicting ICU mortality is summarized in Table 2. XGBoost achieved the highest AUROC (0.93), followed closely by Random Forest (0.91). The DNN model performed moderately well with an AUROC of 0.89, while Logistic Regression had the lowest performance with an AUROC of 0.82. These results indicate that ensemble methods (XGBoost and Random Forest) are particularly effective for mortality prediction due to their ability to model complex data patterns.

Table 2. AUROC of Different Models for Mortality Prediction

Chart 1 illustrates the comparative performance of the four models, clearly showing the superiority of XGBoost.

Графикон 1. Упоредне перформансе испитиваних модела **Chart 1. Comparative performance of the tested models**



Значај карактеристика

Да бисмо разумели које варијабле најзначајније утичу на предвиђање mortalитета, анализирали смо значај карактеристика користећи XGBoost модел, који је показао најбоље перформансе. Као што је приказано у табели 3, старост, SOFA скор и ниво лактата појавили су се као три главна предиктора. Ови налази су у складу са клиничким очекивањима, јер су већа старост и високи SOFA резултати добро утврђени маркери лоше прогнозе [1, 2].

Табела 3. Десет најважнијих карактеристика које утичу на mortalитет на интензивној нези (XGBoost модел)

Feature Importance

To understand which variables most significantly impact mortality prediction, we analysed feature importance using the XGBoost model, which demonstrated the best performance. As presented in Table 3, age, SOFA score, and lactate level emerged as the top three predictors. These findings align with clinical expectations, as older age and high SOFA scores are well-established markers of poor prognosis [1, 2].

Table 3. Top 10 Features Influencing ICU Mortality (XGBoost Model)

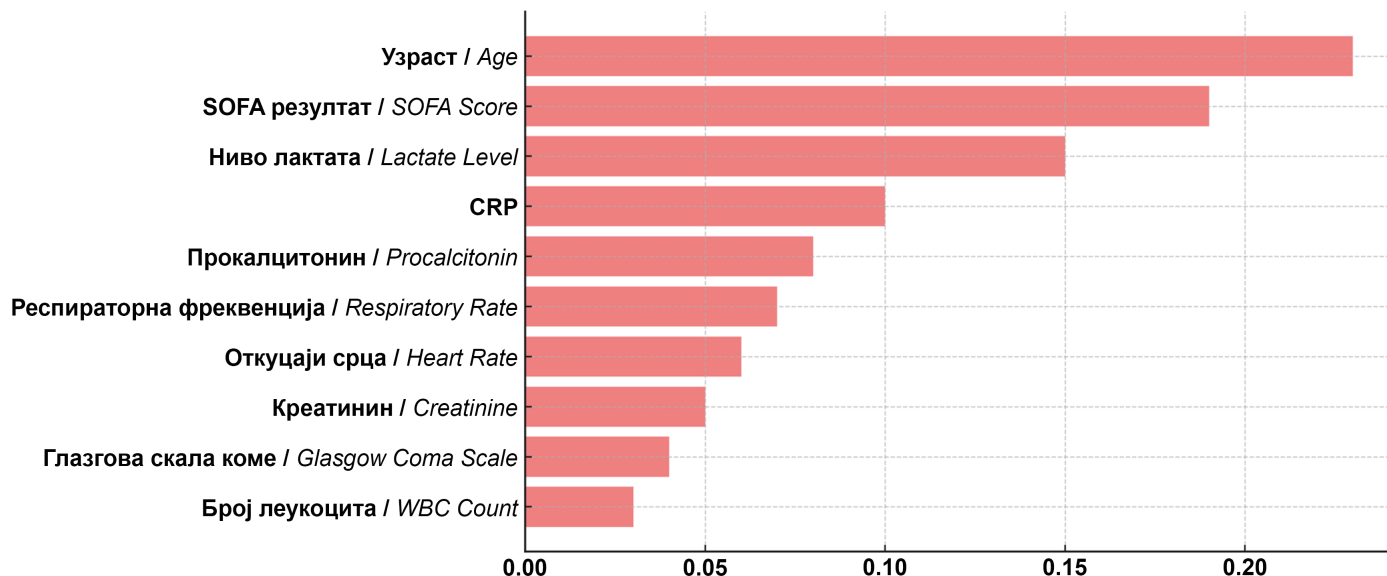
Карактеристика / Feature	Релативна важност / Relative Importance
Узраст / Age	0,23
SOFA резултат / SOFA Score	0,19
Ниво лактата / Lactate Level	0,15
CRP / CRP	0,10
Прокалцитонин / Procalcitonin	0,08
Респираторна фреквенција / Respiratory Rate	0,07
Откуцаји срца / Heart Rate	0,06
Креатинин / Creatinine	0,05
Глазгова скала коме / Glasgow Coma Scale	0,04
Број леукоцита / WBC Count	0,03

Графикон 2 визуелно приказује релативни значај ових карактеристика, наглашавајући доминантну улогу старости и SOFA резултати.

Chart 2 visually depicts the relative importance of these features, emphasizing the dominant role of age and severity scores.

Графикон 2. Релативни значај предиктивних карактеристика за морталитет код пацијената на интензивној нези.

Chart 2. Relative importance of mortality predictors in ICU patients



Предвиђање дужине боравка

Length of Stay Prediction

Тачност предвиђања LOS-а у различитим моделима приказана је у табели 4. XGBoost и Random Forest су поново показали супериорне перформансе, са MAE вредностима од 3,2 и 3,4 дана, респективно. DNN модели су били мање поуздани (MAE 4,0 дана), док је Логистичка регресија била најмање тачна (MAE 4,8 дана).

The accuracy of LOS predictions across different models is shown in Table 4. XGBoost and Random Forest again showed superior performance, with MAE values of 3.2 and 3.4 days, respectively. DNN models were less reliable (MAE 4.0 days), while Logistic Regression was the least accurate (MAE 4.8 days).

Табела 4. Перформансе модела у предвиђању дужине боравка (MAE)

Table 4. Model Performance in Predicting Length of Stay (MAE)

Модел / Model	MAE (дани) / MAE (days)
XGBoost	3,2
Случајна шума / Random Forest	3,4
Дубока неуронска мрежа / Deep Neural Network	4,0
Логистичка регресија / Logistic Regression	4,8

Предвиђање терапијског одговора

Therapeutic Response Prediction

Способност модела да предвиде терапијски одговор у првих седам дана од пријема на интензивну негу приказана је у табели 5. XGBoost је постигао највиши F1-скор од 0,81, затим следе Random Forest (0,79) и DNN (0,77). Логистичка регресија је показала најгоре резултате, са F1-скором од 0,70.

The models' ability to predict therapeutic response within the first seven days of ICU admission is presented in Table 5. XGBoost achieved the highest F1-score of 0.81, followed by Random Forest (0.79) and DNN (0.77). Logistic Regression performed the worst, with an F1-score of 0.70.

Табела 5. F1-резултати за предвиђање терапијског одговора

Table 5. F1-Scores for Therapeutic Response Prediction

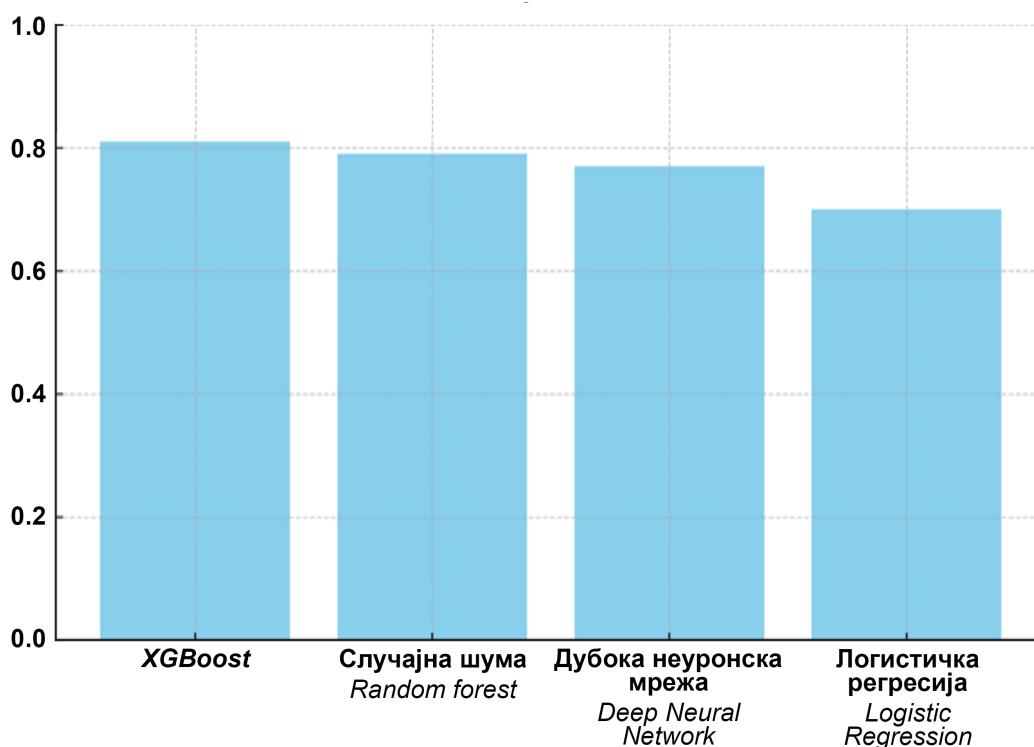
Модел / Model	F1-Резултат / F1-Score
XGBoost	0,81
Случајна шума / Random Forest	0,79
Дубока неуронска мрежа / Deep Neural Network	0,77
Логистичка регресија / Logistic Regression	0,70

Графикон 3 упоређује F1 резултате међу моделима, демонстрирајући XGBoost-ову конзистентну предност.

Chart 3 compares the F1-scores across the models, demonstrating XGBoost's consistent advantage.

Графикон 3. Способност модела за предвиђање терапијског одговора приказана као F1 скор

Chart 3. Models' ability to predict therapeutic response shown as F1 score



Тумачење

Ови резултати потврђују супериорну предиктивну способност ансамбл метода, посебно XGBoost-а, који ефикасно интегрише демографске, клиничке и лабораторијске податке. Високе перформансе у предвиђању mortalитета и терапијског одговора наглашавају њен потенцијал за подршку клиничком одлучивању у реалном времену. Ниже перформансе логистичке регресије одражавају њена ограничења у обухватању сложених интеракција типичних за податке прикупљене из јединица интензивне неге.

Дискусија

Ова студија је истраживала примену модела вештачке интелигенције (ВИ) за предвиђање исхода код паци-

Interpretation

These results confirm the superior predictive capability of ensemble methods, particularly XGBoost, which effectively integrates demographic, clinical, and laboratory data. The high performance in both mortality and therapeutic response prediction underscores its potential for real-time clinical decision support. Logistic Regression's lower performance reflects its limitations in capturing complex interactions typical of ICU data.

Discussion

This study investigated the application of artificial intelligence (AI) models to predict outcomes in patients with severe bacterial infections admitted to intensive care units (ICUs). The hypothesis was that AI models, particularly

јената са тешким бактеријским инфекцијама примљених на одељења интензивне неге (ОИН). Хипотеза је била да би модели вештачке интелигенције, посебно методе ансамбла као што су *XGBoost* и *Random Forest*, надмашили традиционалне системе бодовања попут APACHE II и SOFA у предвиђању морталитета, дужине боравка (LOS) и терапијског одговора. Резултати ове студије снажно подржавају ову хипотезу, показујући да модели вештачке интелигенције заиста нуде супериорну предиктивну тачност у поређењу са конвенционалним методама.

Предвиђање морталитета

Примарни циљ ове студије био је предвиђање морталитета код критично болесних пацијената са бактеријским инфекцијама. Међу тестираним моделима вештачке интелигенције, *XGBoost* је показао највећу тачност, постигавши AUROC од 0,93. Случајна шума је одмах следила са AUROC од 0,91, док је дубока неуронска мрежа (DNN) постигла 0,89. Логистичка регресија, која се често користи као основа, показала је најниже перформансе са AUROC од 0,82. Ови налази указују на то да су методе ансамбла посебно ефикасне за предвиђање морталитета у окружењима интензивне неге, вероватно због њихове способности да обухвате сложене и нелинеарне односе између варијабли. Супериорне перформансе *XGBoost*-а могу се приписати његовој робусности у обрађивању хетерогених података и његовој способности да ефикасно управља интеракцијама карактеристика [1, 2].

Наши налази се поклапају са недавним студијама које су показале ефикасност машинског учења у предвиђању морталитета међу пацијентима на интензивној нези. На пример, Лим и др. су известили да је модел машинског учења ансамбла значајно надмашио традиционалне системе бодовања у предвиђању краткорочног морталитета, са AUROC-ом упоредивим са оним постигнутим у нашој студији [3]. Слично томе, друга истраживања су показала да укључивање клиничких података у реалном времену у моделе машинског учења побољшава предиктивну тачност код критично оболелих популација.

Значај предиктивних карактеристика

Анализа важности карактеристика показала је да су старост, SOFA скор и ниво лактата били најутицајнији предиктори морталитета у *XGBoost* моделу. Ове варијабле су у складу са клиничким очекивањима, јер су већа старост и већи скорови тежине болести дуго повезани са повећаним ризиком од морталитета код пацијената на интензивној нези [1, 2]. Значај инфламаторних маркера, као што су Ц-реактивни протеин (CRP) и про-

ensemble methods such as *XGBoost* and *Random Forest*, would outperform traditional scoring systems like APACHE II and SOFA in predicting mortality, length of stay (LOS), and therapeutic response. The results of this study strongly support this hypothesis, demonstrating that AI models indeed offer superior predictive accuracy compared to conventional methods.

Predicting Mortality

The primary goal of this study was to predict mortality in critically ill patients with bacterial infections. Among the AI models tested, *XGBoost* demonstrated the highest accuracy, achieving an AUROC of 0.93. *Random Forest* followed closely with an AUROC of 0.91, while the *Deep Neural Network (DNN)* achieved 0.89. *Logistic Regression*, often used as a baseline, showed the lowest performance with an AUROC of 0.82. These findings indicate that ensemble methods are particularly effective for mortality prediction in ICU settings, likely due to their ability to capture complex and nonlinear relationships between variables. The superior performance of *XGBoost* can be attributed to its robustness in handling heterogeneous data and its ability to manage feature interactions effectively [1, 2].

Our findings align with recent studies that have demonstrated the effectiveness of machine learning in predicting mortality among ICU patients. For instance, Lim et al. reported that an ensemble machine learning model significantly outperformed traditional scoring systems in predicting short-term mortality, with an AUROC comparable to that achieved in our study [3]. Similarly, other research has shown that incorporating real-time clinical data into machine learning models enhances predictive accuracy in critically ill populations.

Importance of Predictive Features

Feature importance analysis revealed that age, SOFA score, and lactate level were the most influential predictors of mortality in the *XGBoost* model. These variables are consistent with clinical expectations, as older age and higher severity scores have long been associated with increased mortality risk in ICU patients [1, 2]. The importance of inflammatory markers, such as C-reactive protein (CRP) and procalcitonin, also aligns with their known role in indicating severe systemic infection. Additionally, respiratory rate, heart rate, and renal function markers (like creatinine) contributed moderately to the model's predictive accuracy. This finding highlights the multidimensional nature of predicting critical outcomes, where both demographic and physiological parameters play significant roles [4].

калцитонин, такође се поклапа са њиховом познатом улогом у указивању на тешку системску инфекцију. Поред тога, фреквенција дисања, срчани ритам и маркери бубрежне функције (попут креатинина) умерено су допринели предиктивној тачности модела. Ово откриће истиче вишедимензионалну природу предвиђања критичних исхода, где и демографски и физиолошки параметри играју значајну улогу [4].

Предвиђање дужине боравка

Прецизно предвиђање дужине боравка на интензивној нези (LOS) је кључно за управљање ресурсима и планирање неге пацијената. Модели *XGBoost* и *Random Forest* показали су најбоље перформансе у предвиђању LOS-а, са средњим апсолутним грешкама (MAE) од 3,2 и 3,4 дана, респективно. DNN модели су били умерено тачни (MAE 4,0 дана), док је логистичка регресија заостајала (MAE 4,8 дана). Супериорне перформансе метода заснованих на стаблу у предвиђању LOS-а могу се приписати њиховој способности да моделирају сложене интеракције између клиничких варијабли и карактеристика пацијената [5]. Ови резултати сугеришу да предвиђања LOS-а вођена вештачком интелигенцијом могу помоћи у оптимизацији искоришћености кревета на интензивној нези и планирању отпуста. Претходне студије су такође истакле корисност модела вештачке интелигенције у предвиђању LOS-а, посебно када се укључују динамички клинички параметри.

Предвиђање терапијског одговора

Предвиђање терапијског одговора у првих седам дана од пријема на интензивну негу био је још један важан аспект ове студије. *XGBoost* је постигао највиши F1-скор од 0,81, затим су следили *Random Forest* (0,79) и DNN (0,77). Логистичка регресија, са F1-скором од 0,70, поново се показала мање ефикасном. Виши F1-скорови *XGBoost*-а и *Random Forest*-а указују на то да су ови модели бољи у разликовању пацијената који ће вероватно реаговати на лечење од оних који неће. Рана идентификација пацијената са повољном прогнозом омогућава усмереније и ефикасније терапијске интервенције, потенцијално смањујући морбидитет и морталитет на интензивној нези [6].

Поређење са традиционалним системима бодовања

Традиционални системи бодовања попут APACHE II и SOFA се широко користе за прогнозирање на одељењима интензивне неге. Међутим, AUROC-ови ових система (0,85 и 0,87, респективно) били су значајно нижи од оних код тестираних AI модела. Ово откриће истиче ограничења статичких система бодовања, који можда не обухватају у потпуности динамичку и мултифак-

Predicting Length of Stay

Accurately predicting the length of ICU stay (LOS) is crucial for resource management and planning patient care. The XGBoost and Random Forest models demonstrated the best performance in LOS prediction, with mean absolute errors (MAE) of 3.2 and 3.4 days, respectively. DNN models were moderately accurate (MAE 4.0 days), while Logistic Regression lagged behind (MAE 4.8 days). The superior performance of tree-based methods in LOS prediction can be attributed to their ability to model complex interactions between clinical variables and patient characteristics [5]. These results suggest that AI-driven predictions of LOS can aid in optimizing ICU bed utilization and discharge planning. Previous studies have also highlighted the utility of AI models in predicting LOS, particularly when incorporating dynamic clinical parameters.

Predicting Therapeutic Response

Therapeutic response prediction within the first seven days of ICU admission was another important aspect of this study. XGBoost achieved the highest F1-score of 0.81, followed by Random Forest (0.79) and DNN (0.77). Logistic Regression, with an F1-score of 0.70, again proved less effective. The higher F1-scores of XGBoost and Random Forest indicate that these models are better at distinguishing between patients likely to respond to treatment and those who are not. Early identification of patients with a favourable prognosis allows for more targeted and efficient therapeutic interventions, potentially reducing ICU morbidity and mortality [6].

Comparison with Traditional Scoring Systems

Traditional scoring systems like APACHE II and SOFA have been widely used for prognostication in ICUs. However, the AUROCs of these systems (0.85 and 0.87, respectively) were significantly lower than those of the AI models tested. This finding highlights the limitations of static scoring systems, which may not fully capture the dynamic and multifactorial nature of critical illness. AI models, on the other hand, benefit from the integration of diverse data types and real-time updates, offering more precise and individualized predictions [7].

Clinical Implications

The implementation of AI models, particularly ensemble methods, in ICU practice could revolutionize clinical decision-making by enabling more accurate predictions of mortality, LOS, and therapeutic outcomes. This could facilitate early identification of high-risk patients and allow for the optimization of treatment plans, ultimately improving patient care. However, integrating these models into routine

торску природу критичних болести. AI модели, с друге стране, имају користи од интеграције различитих типова података и ажурирања у реалном времену, нудећи прецизности и индивидуализованија предвиђања [7].

Клиничке импликације

Имплементација модела вештачке интелигенције (AI), посебно ансамбл метода, у пракси јединице интензивне неге могла би револуционисати клиничко доношење одлука омогућавајући прецизније предвиђања морталитета, локалне грознице и терапијских исхода. Ово би могло олакшати рану идентификацију пацијената са високим ризиком и омогућити оптимизацију планова лечења, што би на крају побољшало негу пацијената. Међутим, интеграција ових модела у рутинску праксу захтева решавање изазова везаних за интеграцију података, интерпретабилност модела и прихватање од стране лекара. Развој корисничких интерфејса и обезбеђивање транспарентности модела биће кључни за успешно усвајање.

Ограничења

Упркос обећавајућим резултатима, ова студија има неколико ограничења. Прво, ретроспективна природа анализе може увести пристрасности у вези са квалитетом и потпуношћу података. Поред тога, студија се ослањала на податке из више центара интензивне неге, што, иако повећава могућност генерализације, може такође увести варијабилност у клиничку праксу. Потребне су даље проспективне студије како би се потврдили ови налази и истражила имплементација модела предвиђања заснованих на вештачкој интелигенцији у реалним условима интензивне неге [7].

Будући правци

Будућа истраживања требало би да се фокусирају на развој интерпретабилних модела вештачке интелигенције који се могу беспрекорно интегрисати у клиничке токове рада. Штавише, проспективна валидација у различитим здравственим установама је неопходна како би се осигурала робусност и генерализација модела. Поред тога, истраживање модела специфичних за пацијента који се прилагођавају индивидуалним профилима ризика могло би додатно побољшати тачност предвиђања и клиничку корисност.

Ова студија показује да модели вештачке интелигенције, посебно *XGBoost*, значајно надмашују традиционалне системе бодовања у предвиђању морталитета, локалне инфекције и терапијског одговора на интензивној нези код пацијената са тешким бактеријским инфекцијама. Интеграција таквих модела у клиничку праксу има

practice requires addressing challenges related to data integration, model interpretability, and clinician acceptance. Developing user-friendly interfaces and ensuring model transparency will be crucial for successful adoption.

Limitations

Despite the promising results, this study has several limitations. Firstly, the retrospective nature of the analysis may introduce biases related to data quality and completeness. Additionally, the study relied on data from multiple ICU centres, which, while increasing generalizability, may also introduce variability in clinical practices. Further prospective studies are necessary to validate these findings and explore the implementation of AI-based prediction models in real-world ICU settings [7].

Future Directions

Future research should focus on developing interpretable AI models that can be seamlessly integrated into clinical workflows. Moreover, prospective validation across diverse healthcare settings is essential to ensure the robustness and generalizability of the models. Additionally, investigating patient-specific models that adapt to individual risk profiles could further enhance prediction accuracy and clinical utility.

This study demonstrates that AI models, particularly XGBoost, significantly outperform traditional scoring systems in predicting ICU mortality, LOS, and therapeutic response in patients with severe bacterial infections. The integration of such models into clinical practice holds the potential to improve patient outcomes through early risk stratification and personalized therapeutic planning. Continued research is needed to address the challenges of model implementation and validation in diverse ICU environments.

Conclusion

This study demonstrates the potential of artificial intelligence (AI) models, particularly ensemble methods like XGBoost and Random Forest, in predicting outcomes among patients with severe bacterial infections admitted to intensive care units (ICUs). The primary objective was to assess whether AI-based models could outperform traditional scoring systems such as APACHE II and SOFA in predicting mortality, length of stay (LOS), and therapeutic response. The results strongly support this hypothesis, as AI models consistently showed superior predictive performance.

Key Findings

The XGBoost model demonstrated the highest accuracy

потенцијал да побољша исходе пацијената кроз рану стратификацију ризика и персонализовано терапијско планирање. Потребна су континуирана истраживања како би се решили изазови имплементације и валидације модела у различитим окружењима интензивне неге.

Закључак

Ова студија демонстрира потенцијал модела вештачке интелигенције (ВИ), посебно ансамбл метода попут *XGBoost* и *Random Forest*, у предвиђању исхода код пацијената са тешким бактеријским инфекцијама примљених на одељења интензивне неге (ОИН). Примарни циљ је био да се процени да ли модели засновани на ВИ могу да надмаше традиционалне системе бодовања као што су APACHE II и SOFA у предвиђању морталитета, дужине боравка (LOS) и терапијског одговора. Резултати снажно подржавају ову хипотезу, јер су модели ВИ константно показивали супериорне предиктивне перформансе.

Кључни налази

Модел *XGBoost* је показао највећу тачност у предвиђању морталитета, постигавши AUROC од 0,93, значајно надмашујући конвенционалне системе бодовања. Ова предност се приписује способности модела да интегрише различите клиничке и лабораторијске варијабле, хватајући сложене интеракције које традиционалне методе често превиђају. Модел *Random Forest* је такође показао робусне перформансе са AUROC од 0,91, истичући ефикасност метода ансамбла заснованих на дрвету. Насупрот томе, логистичка регресија, која се често користи као основа, показала је најнижу тачност, што одражава њен ограничени капацитет за моделирање нелинеарних односа својствених подацима са интензивне неге.

Анализа важности карактеристика показала је да су старост, SOFA скор и ниво лактата били најутицајнији предиктори морталитета. Ови налази се поклапају са утврђеним клиничким знањем, јер су већа старост и већи скорови тежине болести добро документовани фактори ризика у интензивној нези. Поред тога, маркери инфламације попут CRP и прокалцитонина идентификовани су као значајни предиктори, што наглашава улогу системске упале у одређивању исхода пацијената.

Што се тиче предвиђања LOS-а, *XGBoost* и *Random Forest* модели су поново показали супериорну тачност у поређењу са DNN-ом и Логистичком регресијом, са средњим апсолутним грешкама (MAE) од 3,2 и 3,4 дана, респективно. Ово сугерише да су ансамбл методе посебно веште у хватању варијабилности повезане

cy for mortality prediction, achieving an AUROC of 0.93, significantly outperforming conventional scoring systems. This advantage is attributed to the model's ability to integrate diverse clinical and laboratory variables, capturing complex interactions that traditional methods often overlook. Random Forest also showed robust performance with an AUROC of 0.91, highlighting the efficacy of tree-based ensemble methods. In contrast, Logistic Regression, often used as a baseline, showed the lowest accuracy, reflecting its limited capacity to model nonlinear relationships inherent in ICU data.

Feature importance analysis revealed that age, SOFA score, and lactate level were the most influential predictors of mortality. These findings align with established clinical knowledge, as older age and higher severity scores are well-documented risk factors in critical care. Additionally, inflammatory markers such as CRP and procalcitonin were identified as significant predictors, underscoring the role of systemic inflammation in determining patient outcomes.

In terms of predicting LOS, XGBoost and Random Forest models again demonstrated superior accuracy compared to DNN and Logistic Regression, with mean absolute errors (MAE) of 3.2 and 3.4 days, respectively. This suggests that ensemble methods are particularly adept at capturing the variability associated with recovery trajectories in critically ill patients. Furthermore, the prediction of therapeutic response within the first seven days showed that XGBoost achieved the highest F1-score (0.81), reinforcing the model's utility in identifying patients likely to benefit from early interventions.

Clinical Implications

The integration of AI-based predictive models into ICU practice could substantially enhance clinical decision-making. By accurately identifying high-risk patients and predicting therapeutic responses, healthcare providers can tailor interventions more effectively, potentially reducing morbidity and mortality. Additionally, reliable LOS predictions could improve resource allocation, optimizing ICU bed utilization and discharge planning. However, translating these models into routine practice requires addressing challenges related to data integration, model interpretability, and clinician acceptance. Developing user-friendly interfaces and providing education on AI model usage are essential steps to ensure successful implementation.

Limitations and Future Directions

Despite the promising results, several limitations must be acknowledged. The study's retrospective design introduces inherent biases related to data quality and completeness. Moreover, the variability between ICU centres might affect

са путањама опоравка код критично болесних пацијената. Штавише, предвиђање терапијског одговора у првих седам дана показало је да је *XGBoost* постигао највиши F1-скор (0,81), појачавајући корисност модела у идентификацији пацијената који ће вероватно имати користи од раних интервенција.

Клиничке импликације

Интеграција предиктивних модела заснованих на вештачкој интелигенцији у праксу одељења интензивне неге могла би значајно побољшати клиничко доношење одлука. Прецизним идентификовањем пацијената високог ризика и предвиђањем терапијских одговора, здравствени радници могу ефикасније прилагодити интервенције, потенцијално смањујући морбидитет и морталитет. Поред тога, поуздана предвиђања LOS-а могла би побољшати расподелу ресурса, оптимизујући искоришћеност кревета на интензивној нези и планирање отпуста. Међутим, превођење ових модела у рутинску праксу захтева решавање изазова везаних за интеграцију података, интерпретабилност модела и прихватање од стране лекара. Развој корисничких интерфејса и пружање едукације о коришћењу модела вештачке интелигенције су неопходни кораци за осигуравање успешне имплементације.

Ограничења и будући правци

Упркос обећавајућим резултатима, мора се признати неколико ограничења. Ретроспективни дизајн студије уводи инхерентне пристрасности везане за квалитет и потпуност података. Штавише, варијабилност између центара интензивне неге може утицати на генерализацију модела. Потребне су проспективне мултицентричне студије како би се потврдили ови налази и проценила применљивост у стварном свету. Поред тога, развој интерпретабилнијих модела вештачке интелигенције који се беспрекорно интегришу са постојећим електронским здравственим картонима (ЕЗК) биће кључан за стицање поверења лекара и подстицање широког усвајања.

Будућа истраживања требало би да се фокусирају на побољшање транспарентности модела и истраживање адаптивних алгоритама специфичних за пацијента, који се континуирано ажурирају како нови клинички подаци постају доступни. Укључивање токова података у реалном времену у системе за подршку одлучивању вођене вештачком интелигенцијом могло би додатно повећати њихову клиничку корисност. Штавише, истраживање хибридних модела који комбинују машинско учење са традиционалним системима бодовања може дати још тачнија предвиђања.

the generalizability of the models. Prospective multicentre studies are needed to validate these findings and assess real-world applicability. Additionally, developing more interpretable AI models that integrate seamlessly with existing electronic health records (EHRs) will be crucial for gaining clinician trust and fostering widespread adoption.

Future research should focus on enhancing model transparency and investigating patient-specific adaptive algorithms that continuously update as new clinical data become available. Incorporating real-time data streams into AI-driven decision support systems could further increase their clinical utility. Moreover, exploring hybrid models that combine machine learning with traditional scoring systems may yield even more accurate predictions.

Final Remarks

This study highlights the significant potential of AI-driven predictive models to improve patient outcomes in critical care settings. By leveraging the strengths of ensemble learning methods, healthcare professionals can make more informed decisions, leading to better patient management and resource optimization. As AI continues to evolve, its integration into intensive care practices holds the promise of transforming patient care, but it must be approached with careful validation and consideration of ethical and practical challenges.

Завршне напомене

Ова студија истиче значајан потенцијал предиктивних модела вођених вештачком интелигенцијом за побољшање исхода пацијената у установама интензивне неге. Коришћењем предности метода ансамбл учења, здравствени радници могу доносити информисаније одлуке, што доводи до бољег управљања пацијентима и оптимизације ресурса. Како се вештачка интелигенција наставља развијати, њена интеграција у праксе интензивне неге обећава трансформацију неге пацијената, али јој се мора приступити уз пажљиву валидацију и разматрање етичких и практичних изазова.

Литература / References

1. Gao J, Lu Y, Ashrafi N, Domingo I, Alaei K, Pishgar M. Prediction of sepsis mortality in ICU patients using machine learning methods. *BMC Med Inform Decis Mak.* 2024; 24:228. <https://doi.org/10.1186/s12911-024-02630-z>
2. Goldschmidt E, Rannon E, Bernstein D, Wasserman A, Roimi M, Shrot A, et al. Predicting appropriateness of antibiotic treatment among ICU patients with hospital-acquired infection. *NPJ Digit Med.* 2025; 8:87. <https://doi.org/10.1038/s41746-024-01426-9>
3. Lim L, Gim U, Cho K, Yoo D, Ryu HG, Lee HC. Real-time machine learning model to predict short-term mortality in critically ill patients: development and international validation. *Crit Care.* 2024; 28:76. <https://doi.org/10.1186/s13054-024-04866-7>
4. Rocheteau E, Liò P, Hyland S. Temporal Pointwise Convolutional Networks for Length of Stay Prediction in the Intensive Care Unit. *arXiv:2007.09483 [Preprint].* 2020 [cited 2026 Feb 8]. <https://doi.org/10.48550/arXiv.2007.09483>
5. Nallabasannagari AR, Reddiboina M, Seltzer R, Zeffiro T, Sharma A, Bhandari M. All Data Inclusive, Deep Learning Models to Predict Critical Events in the Medical Information Mart for Intensive Care III Database (MIMIC III). *arXiv:2009.01366 [Preprint].* 2020 [cited 2026 Feb 8]. <https://doi.org/10.48550/arXiv.2009.01366>
6. Zhang J, Bolanos L, Tanwar A, Ive J, Gupta V, Guo Y. Clinical Utility of the Automatic Phenotype Annotation in Unstructured Clinical Notes: ICU Use Cases. *arXiv:2107.11665 [Preprint].* 2021 [cited 2026 Feb 8]. <https://doi.org/10.48550/arXiv.2107.11665>
7. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, Motaei A, Madkour M, Pardalos PM, et al. MySurgeryRisk: Development and Validation of a Machine-learning Risk Algorithm for Major Complications and Death After Surgery. *Ann. Surg.* 2019; 269(4):652–62. <https://doi.org/10.1097/SLA.0000000000002706>



Примљено / Received
8. 2. 2026.

Ревидирано / Revised
12. 6. 2026.

Прихваћено / Accepted
12. 6. 2026.

Кореспонденција / Correspondence

Шћепан Синановић – Šćepan Sinanović
scep.an.sinanovic@gmail.com

ORCID

Šćepan Sinanović
<https://orcid.org/0000-0002-8125-7873>
Marko Kimi Milić
<https://orcid.org/0009-0007-1308-7220>