ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.57431
Nam Kem Med J 2025;13(3):209-221

Naltrexone Induced Agitation Management: Employing a Hybrid Artificial Neural Network Model to Determine the Appropriate Dosage of Intravenous Diazepam

Naltrekson Kaynaklı Ajitasyon Yönetimi: İntravenöz Diazepamın Uygun Dozajını Belirlemek için Hibrit Yapay Sinir Ağı Modeli Kullanılması

Seyed Ali MOHTARAMI[1], Sayed Masoud HOSSEINI[2], Shahin SHADNIA[2], Babak MOSTAFAZADEH[2], Peyman Erfan Talab EVINI[2], Omid MEHRPOUR[3], Shakiba FARHADPUR[4], Mitra RAHIMI[2]

[1]Payame Noor University (PNU) Faculty of Engineering, Department of Computer Engineering and Information Technology, Tehran, Iran
[2]Toxicological Research Center, Excellence Center of Clinical Toxicology, Department of Clinical Toxicology, Loghman Hakim Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3]Michigan Poison & Drug Information Center, Wayne State University Faculty of Medicine, Michigan, USA
[4]Department of Pharmacology, School of Pharmacy, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ABSTRACT

Aim: Agitation is a prevalent symptom of opioid withdrawal caused by naltrexone. Managing agitation induced by naltrexone poses challenges as current drug interventions are either ineffective or require intensive care. This study sought to determine the most effective diazepam dosage for alleviating naltrexone-induced agitation.

Materials and Methods: This research examined a dataset comprising 615 patient medical records from Loghman Hakim Hospital in Tehran, Iran, focusing on cases of patients experiencing naltrexone-induced agitation. The dataset included individuals who were administered low-dose diazepam (<10 mg; 383 cases) and high-dose diazepam (>10 mg; 232 cases). The predictive performance of the developed models was assessed based on metrics such as accuracy, specificity, sensitivity, F1-score, and ROC curve analysis.

Results: The bat algorithm demonstrated the highest performance among meta-heuristic algorithms, achieving a score of 89.5% (0.895) at iteration 128. A comparative evaluation of five decision tree classifiers revealed that the Extra Trees Classifier surpassed others, attaining an accuracy of 0.8649, sensitivity of 0.8649, precision of 0.8645, F1-score of 0.8649, and area under the curve (AUC) of 0.9343. Following the determination of feature importance and training of a multilayer perceptron neural network with weighted features, the model exhibited superior performance with an accuracy of 0.91, sensitivity of 0.9, precision of 0.92, F1-score of 0.91, and AUC of 0.94.

Conclusion: Features for predicting the appropriate dose of diazepam in patients with naltrexone-induced agitation included recent opioid use, Richmond Agitation-Sedation scale, amount of ingested naltrexone, pulse rate, systolic blood pressure, level of consciousness, serum levels of sodium, creatinine, and lactate dehydrogenase. Our research findings indicate that a weighted multilayer perceptron neural network shows promise in accurately forecasting the necessity of increased doses of diazepam for patients experiencing naltrexone-induced agitation. This is particularly evident when utilizing meta-heuristic techniques for feature selection and assigning importance of selected features based on the classifier with the highest AUC. This model could guide clinicians in tailoring diazepam doses to manage naltrexone induced agitation safely.

Keywords: Naltrexone, agitation, diazepam, meta-heuristic algorithm, multilayer perceptron, machine learning

Address for Correspondence: Mitra RAHIMI MD, Toxicological Research Center, Excellence Center of Clinical Toxicology, Department of Clinical Toxicology, Loghman Hakim Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
E-mail: mrahimi744@gmail.com **ORCID ID:** orcid.org/0000-0002-8785-8207
Received: 07.12.2024 **Accepted:** 02.05.2025 **Publication Date:** 07.10.2025

Cite this article as: Rahimi M, Mohtarami SA, Hosseini SM, Shadnia S, Farhadpour S, Mehrpour O, et al. Naltrexone induced agitation management: employing a hybrid artificial neural network model to determine the appropriate dosage of intravenous diazepam. Nam Kem Med J. 2025;13(3):209-221

ÖZ

Amaç: Ajitasyon, naltreksonun neden olduğu opioid yoksunluğunun yaygın bir semptomudur. Naltreksonun neden olduğu ajitasyonu yönetmek, mevcut ilaç müdahalelerinin ya etkisiz olması ya da yoğun bakım gerektirmesi nedeniyle zorluklar ortaya çıkarmaktadır. Bu çalışma, naltreksonun neden olduğu ajitasyonu hafifletmek için en etkili diazepam dozajını belirlemeyi amaçlamaktadır.

Gereç ve Yöntem: Bu araştırma, İran'ın Tahran kentindeki Loghman Hakim Hastanesi'nden alınan 615 hasta tıbbi kayıtlarından oluşan bir veri setini incelemiş ve naltreksonun neden olduğu ajitasyon olgularına odaklanmıştır. Veri seti, düşük doz diazepam (<10 mg; 383 olgu) ve yüksek doz diazepam (>10 mg; 232 olgu) verilen bireyleri içermektedir. Geliştirilen modellerin tahmin performansı, doğruluk, özgüllük, duyarlılık, F1 skoru ve ROC eğrisi analizi gibi metriklere göre değerlendirildi.

Bulgular: Yarasa algoritması, meta-sezgisel algoritmalar arasında en yüksek performansı göstererek, 128. yinelemede %89,5 (0,895) puan aldı. Beş karar ağacı sınıflandırıcısının karşılaştırmalı değerlendirmesi, Ekstra Ağaç Sınıflandırıcı diğerlerini geride bırakarak 0,8649 doğruluk, 0,8649 duyarlılık, 0,8645 kesinlik, 0,8649 F1 skoru ve 0,9343 eğri altındaki alan (AUC) elde ettiğini ortaya koymuştur. Özelliklerin öneminin belirlenmesi ve ağırlıklı özelliklere sahip çok katmanlı bir algılayıcı sinir ağının eğitilmesinin ardından, model 0,91 doğruluk, 0,9 duyarlılık, 0,92 kesinlik, 0,91 F1 skoru ve 0,94 AUC ile üstün performans sergilemiştir. Naltrekson kaynaklı ajitasyon hastalarında uygun diazepam dozunu tahmin etmek için kullanılan özellikler arasında son zamanlarda opioid kullanımı, Richmond Ajitasyon-Sedasyon ölçeği, alınan naltrekson miktarı, nabız hızı, sistolik kan basıncı, bilinç düzeyi, serum sodyum, kreatinin ve laktat dehidrojenaz düzeyleri yer almaktadır.

Sonuç: Araştırma bulgularımız, ağırlıklı çok katmanlı algılayıcı sinir ağının, naltrekson kaynaklı ajitasyon yaşayan hastalar için diazepam dozunun artırılması gerekliliğini doğru bir şekilde tahmin etmede umut vaat ettiğini göstermektedir. Bu, özellikle özellik seçimi için meta-sezgisel teknikler kullanıldığında ve en yüksek AUC'ye sahip sınıflandırıcıya göre seçilen özelliklerin önemine göre atandığında belirgindir. Bu model, klinisyenlere naltrekson kaynaklı ajitasyonu güvenli bir şekilde yönetmek için diazepam dozlarını ayarlamada rehberlik edebilir.

Anahtar Kelimeler: Naltrekson, ajitasyon, diazepam, meta-sezgisel algoritma, çok katmanlı algılayıcı, makine öğrenimi

INTRODUCTION

Naltrexone, an opioid antagonist, is increasingly being prescribed to address opioid use disorder, alcohol use disorder, and chronic pain[1]. This non-selective opioid antagonist is commonly used for maintenance therapy in opioid dependency due to its long-acting nature and high affinity to μ receptors. The primary long-acting metabolite of naltrexone, six-beta-naltrexone, extends the antagonistic effects of naltrexone on narcotic receptors[2].

There are three potential scenarios in which a naltrexone-related fatality may occur: opioid overdose during oral naltrexone treatment as patients may consume high doses of opioids in an attempt to bypass the blockade; opioid overdose following the cessation of naltrexone treatment as individuals resume opioid use after treatment and lose their tolerance to opioids; and toxicity from naltrexone itself, particularly harmful to the liver in doses exceeding five times the recommended safe dose[3,4].

In preclinical investigations involving opioid-naïve animal subjects, naltrexone demonstrated a relatively low level of acute toxicity[5]. It was observed that naltrexone led to a non-significant reduction in respiratory rate and pupillary size, as well as a significant decrease in body temperature among five individuals with a history of addiction. Clinical trials conducted on opioid use disorder revealed that naltrexone effectively counteracted the effects of heroin for a duration of up to 72 hours and that it exhibited no signs of toxicity at doses of up to 200 mg per day. Research on naltrexone in opiate-naive healthy individuals suggested that the substance may possess certain opiate-like characteristics. Following the administration of 50 mg of naltrexone, participants reported experiencing drowsiness, dysphoria, sexual thoughts, penile erection, and an elevation in luteinizing hormone levels[6,7].

The literature suggests that the mortality risk associated with naltrexone is significantly increased in cases of relapse following opioid abstinence, primarily due to opioid toxicity. Patients undergoing naltrexone treatment experience a reduced tolerance to agonist opioids compared to their pre-treatment levels, making them susceptible to potentially fatal overdoses at the end of a dosing interval, after missing a dose, or upon discontinuation of treatment. Attempts to circumvent the opioid blockade can also result in fatal overdoses[8]. Additionally, the use of naltrexone in opioid-dependent individuals can trigger acute and severe withdrawal symptoms, characterized by heightened agitation compared to withdrawal from abstinence[9].

Diazepam has been commonly used for symptomatic relief, though case reports suggest variable effectiveness, often requiring high doses for sedation. One study reported initial limited relief with 10 mg intravenous diazepam, necessitating escalation to 60 mg for effective symptom control[10]. A comparative trial showed that midazolam had a faster onset of action (67 minutes) compared to diazepam (81 minutes), though neither was deemed ideal for rapid agitation management[11]. Furthermore, a review of primary studies reported that the doses of diazepam administered to treat agitation induced by naltrexone ranged from 5 mg to over 40 mg, depending the different naltrexone formulation and severity of symptoms[12].

There is limited empirical evidence supporting treatment recommendations, and consensus among experts is lacking. Traditional symptomatic therapies like antiemetics, clonidine,

benzodiazepines, and titrated doses of an opioid agonist are generally effective in managing opioid withdrawal symptoms[12]. Various medications have been utilized to address agitation induced by naltrexone in addicted patients, but none have demonstrated satisfactory efficacy in controlling agitation upon admission. Benzodiazepines are considered safe sedatives with documented efficacy in similar scenarios, although their use for managing agitation resulting from inappropriate naltrexone use is uncertain. Prompt management of agitation is crucial during severe episodes, as patients may pose a risk to themselves, companions, or medical staff[13]. Therefore, this study utilized a hybrid artificial neural network model to determine the optimal intravenous diazepam dosage for maximizing the efficacy of benzodiazepines in controlling naltrexone-induced agitation.

MATERIALS AND METHODS

Study Design and Setting

This study is a retrospective cross-sectional analysis of medical records of patients experiencing naltrexone-induced withdrawal symptoms at Loghman Hakim Hospital from April 2002 to March 2016. Trained clinical toxicologists documented patients' medical history, treatment trends, and vital signs. Figure 1 summarize the methodology of study. The study received approval from the Ethics Committee of Shahid Beheshti University of Medical Sciences (decision no: IR.SBMU.RETECH.REC.1402.626, date: 01.07.2024). Patient data were de-identified using file numbers to protect confidentiality.

Data Set Description and Participants

The dataset consists of 615 patient records from Loghman Hakim Hospital, focusing on individuals experiencing naltrexone-induced withdrawal symptoms. This hospital serves as a primary referral center for individuals affected by poisoning cases. Among the dataset entries, 232 cases were treated with high-dose diazepam (>10 mg) while 383 cases received low-dose diazepam (≤10 mg). The study included all patients presenting with naltrexone-induced withdrawal symptoms, whether due to intentional or accidental poisoning, at Loghman Hakim Hospital. The exclusion criteria consisted of cases with multiple drug toxicity, severe chronic comorbidities (e.g., cardiovascular diseases, neurological disorders, psychiatric conditions, or seizure disorders), as well as patients with incomplete medical records related to demographic information, vital signs, or paraclinical data at admission.

Benzodiazepines are considered safe sedatives with documented efficacy in similar scenarios, although their use for managing agitation resulting from inappropriate naltrexone use is uncertain[11].

Data Gathering

A comprehensive examination of patient medical records was conducted by a team of six researchers. Data were extracted from the electronic databases of Loghman Hakim Hospital (Sabara and Shafa databases) using a pre-made checklist. The collected information included demographic details such as age, gender, last opioid intake, and the purpose of naltrexone usage. Additionally, vital signs and withdrawal symptoms upon admission were documented. Furthermore, details on the administration of diazepam, blood glucose levels, electrocardiograms (ECGs), venous blood gases (VBG), blood electrolytes, liver and kidney function tests, and the Richmond Agitation-Sedation scale (RASS) were also recorded.

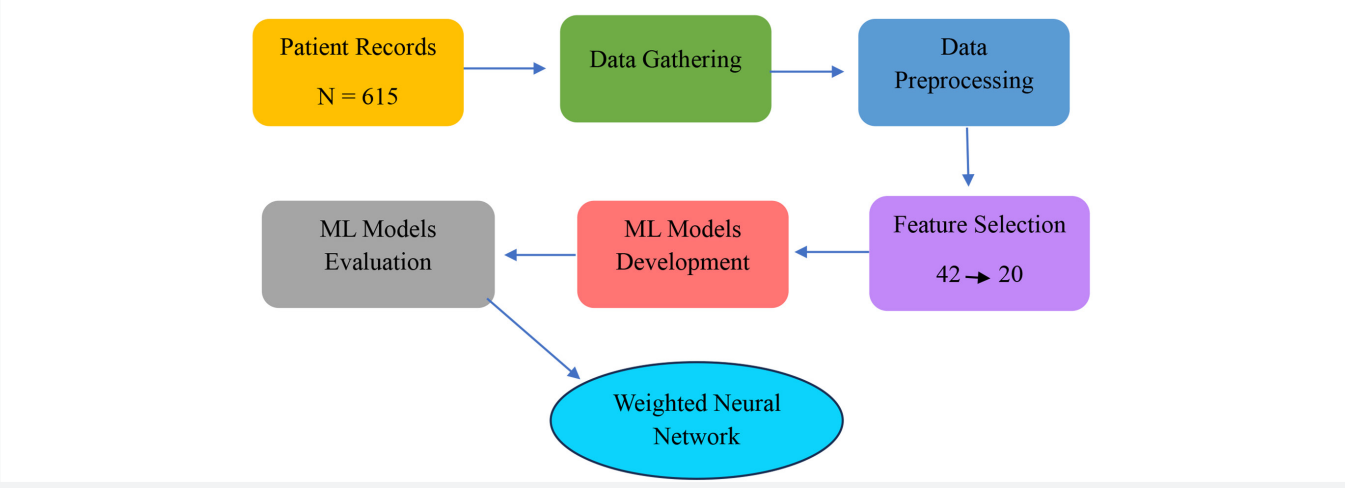

Figure 1. The flowchart visualizing the methodology of study

ML: Machine learning

Pre-processing of the Data

Various preprocessing techniques were used in this study to optimize classification algorithms after data collection. Methods included removing variables with over 70% missing data, identifying and eliminating noisy data and outliers, and addressing missing data in other variables with Mean Imputation and Stochastic Regression Imputation. To mitigate bias from mean imputation, we restricted its use to variables with minimal missingness (<5%). Data points falling outside the normal range were excluded in consultation with clinical committees. The dataset was split into train and test sets, then further divided into Class A for patients receiving diazepam doses of 10 mg or less, and Class B for doses over 10 mg.

Feature Selection

The process of feature selection, which involves eliminating redundant variables from the initial dataset while retaining essential information, is crucial in mitigating the risk of overfitting. In the initial phase of the research, a total of 42 features were identified, encompassing a diverse array of clinical and paraclinical symptoms, reflecting the high complexity and dimensionality of the characteristics under investigation. A crucial aspect of the study involved the selection of an optimal subset of features. To achieve this, the formulation of a loss function was imperative. The loss function was established through the utilization of the following equation in the research process:

$$\text{Loss Function} = \frac{1}{Accuracy}$$

To assess accuracy in the context of medical sciences for predicting treatment methods, where accuracy is crucial, the decision tree family of machine learning (ML) algorithms was employed. This family of algorithms is adept at addressing complex cases and offers valuable insights into feature importance for decision-making processes. Moreover, these algorithms demonstrate resilience when faced with imbalanced datasets. The evaluation of ML models from the decision tree family was conducted based on metrics such as accuracy, precision, sensitivity, area under the curve (AUC), and F1-score.

The decision tree algorithms utilize conditional statements to establish predictive criteria, with each algorithm possessing distinct architectural characteristics beyond this fundamental condition. The dataset was trained using a ten-fold cross-validation technique, involving the partitioning of the data into ten subsets and iteratively applying the holdout method. Meta-parameters were fine-tuned according to the training dataset through the utilization of the cross-validation methodology. The loss function was established based on the predictive accuracy of the model exhibiting the highest AUC metric.

The process of feature selection was conducted through the utilization of metaheuristic algorithms. The bat algorithm is a metaheuristic algorithm that operates on a population-based approach and is designed for addressing continuous optimization problems. This algorithm has demonstrated efficacy in optimizing solutions across various domains such as cloud computing, feature selection, image processing, and control engineering challenges[14].

When employing a search method based on 42 features, the computational complexity increases significantly, as the number of potential states grows exponentially at a rate of 42 squared. Furthermore, this search approach lacks convergence due to its random nature, in contrast to meta-heuristic algorithms which exhibit convergence and reduce the number of potential states as the search space is optimized.

The meta-heuristic algorithm has a time complexity of O (N^2), which is the same as a quadratic polynomial. In order to efficiently address the NP problem, we suggest using binary feature algorithms that have minimal time complexity and cost for organizing and distributing tasks in our feature selection issue. We outline an objective function and provide a table displaying the average duration for each iteration.

In this study, seven meta-heuristic algorithms were employed to select features based on the loss function including Binary Genetic Algorithm, Binary Particle Swarm Optimization, Binary Cuckoo Search, Binary Firefly Algorithm, Binary Bat Algorithm, Binary Gravitational Search Algorithm, and Binary Dragon Fly Algorithm. The meta-heuristic algorithm was chosen for its proven efficiency in medical feature selection[15,16]. The algorithm aims to minimize the loss function and systematically searches for a binary list. The binary list indicates feature selection, with a selection indicator represented by the number one and non-selection indicated by zero. Algorithm with the best score and the least loss function was utilized for feature selection[17].

Statistical Analysis

The Kolmogorov-Smirnov and the Shapiro-Wilk tests results revealed that all continuous variables were distributed non-normally. Consequently, the continuous variables were represented by their median values and interquartile ranges and were analyzed using the Mann-Whitney U test. Categorical variables were reported as absolute frequencies and respective percentage and were analyzed using the chi-square test. The performance of classification models was assessed through the receiver-operating curve. Additionally, other performance metrics including the accuracy, sensitivity, and specificity were computed. In this research, the Python Programming Language (version 13.1) and associated libraries were used. Libraries such as Matplotlib, NumPy, Seaborn, and Pandas were used for data analysis and visualization purposes. The scikit-learn library was employed to develop algorithms and evaluate ML models

performance. For descriptive analyses, the SPSS version 26 was utilized.

Model Evaluation (Stage 2)

After feature selection, five classifiers from the ML realms were employed to construct a predictive model for the appropriate diazepam dosage in patients experiencing naltrexone induced withdrawal. Among the ML models employed were the Light Gradient Boosting Machine, Random Forest Classifier, Gradient Boosting Classifier, Extreme Gradient Boosting and Extra Trees Classifier. The aim of utilizing this array of classifiers was to improve prediction accuracy and gain insights into the intricate factors influencing the optimal diazepam dosage for managing agitation induced by naltrexone. The dataset was divided randomly into training (70%) and testing (30%) sets to develop and validate the ML algorithms. A ten-fold cross-validation technique was applied to train the dataset with 20 selected features, involving the division of the dataset into ten sections and conducting the holdout method iteratively. Hyperparameters were adjusted based on the training dataset using the cross-validation approach. Subsequently, the classification algorithms were tested on the testing dataset to evaluate their performance. The performance of the classifiers in predicting appropriate dosage of diazepam in patients with naltrexone induced agitation was assessed using underfitting and overfitting evaluation methods, along with five standard efficiency testing metrics such as accuracy, specificity, sensitivity, precision, and F1-score according to the following equations:

1) $\text{classification accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} * 100$

2) $\text{classification sensitivity} = \frac{\text{TP}}{\text{TP+FN}} * 100$

3) $\text{classification specificity } = \frac{\text{TN}}{\text{TN+FP}} * 100$

4) $\text{classification percision } = \frac{\text{TP}}{\text{TP + FP}} * 100$

5) $\text{F1} - \text{Score} = \frac{2\times(\text{precision} \times \text{Sensitivity})}{\text{precision+Sensitivity}}$

The performance of each classifier was compared against other ML algorithms using these metrics. The best-performing model was selected based on the efficiency results to proceed with further data analysis and to determine the significance of features for neural network weighting through model tuning.

Feature Weight Calculation Using Decision Trees

In order to ascertain the importance coefficient of 20 selected features, the researchers enlisted the assistance of the most effective classifier based on the AUC metric in the realm of ML. Leveraging the unique characteristic of decision trees in calculating the Gini index, these trees were employed to evaluate feature importance. The efficacy of decision tree algorithms within this family is underscored by the root node, which encompasses all initial data pertinent to the issue at hand, in this instance, the 20 selected features. Subsequently, the attribute selection measure was utilized to identify the optimal features based on their level of importance. The feature that yields the most substantial decrease in impurity within a node is deemed the most valuable. Both Gini and Entropy methodologies can be applied to assess the impurity associated with each attribute. The research utilized the Gini index technique to assess the feature importance. This method involves favoring and choosing features with a lower Gini index over those with a higher Gini index in the decision tree. The Gini index is determined through the following mathematical formula:

$$G(t) = 1 - \sum_{i}^{c} p_i^2$$

The Gini impurity at a given node "t" is denoted as "G (t)", where "pi" represents the proportion of observations belonging to Class C at node "t". The Gini index is determined by subtracting the sum of the squared probabilities of each class from one. The information pertaining to the 20 selected features underwent initial processing through MinMax Scaler. This technique involves scaling the data, ensuring that the minimum feature is set to zero and the maximum feature is set to one. Notably, this approach maintains the original distribution shape of the data.

$$FX_i' = \frac{x_i - x_{min}}{x_{man} - x_{min}}$$

Next, the features intended for incorporation into the neural network were assigned weights based on the subsequent formula:

$$FX_i^{new} = FX_i' * w_i^{\alpha}$$

In the aforementioned relationship, "w" denotes the weight assigned to individual features derived from the decision tree. The parameter α is indicative of the hyperparameter utilized in this context to ascertain the impact of feature weights prior to their integration into the neural network. For the purposes of this research, α was set at a value of one. Finally weighted features were integrated in a multilayer perceptron (MLP) neural network. In this study, we used three layers including the input and output layers and the hidden layer. The weighted neural network's performance was evaluated against the ML models exhibiting the most accurate predictive capabilities[17].

RESULTS

Patient's Characteristics

The medical documentation of 907 individuals presenting with naltrexone toxicity was examined, with 292 patients being excluded based on predetermined exclusion criteria. The patient

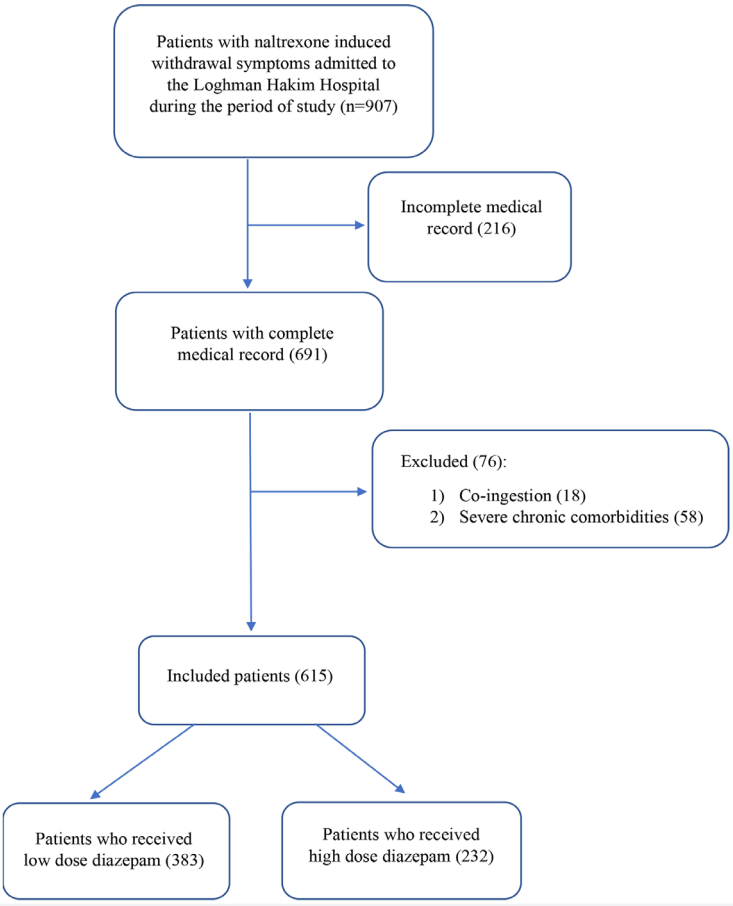

Figure 2. The patient selection flowchart

selection methodology is depicted in Figure 2 for clarity and transparency. The remaining cohort of 615 patients exhibited symptoms of naltrexone-induced agitation, with a mean age of 37.27 years and a standard deviation of 11.52 years, falling within the age range of 14 to 70 years. The study comprised 589 male participants with an average age of 37.27±11.50 and 26 female participants with an average age of 37.23±12.27. Statistical analysis revealed no significant difference in average age between the sexes (p=0.149). Among the 615 cases of poisoning examined, agitation in 232 cases (37.7%) was managed using a high dose of diazepam (exceeding 10 mg), while agitation in 383 cases (62.3%) was controlled with a low dose of diazepam (equal to 10 mg). The descriptive and analytical statistical outcomes pertaining to these two dosage categories are presented in Table 1.

Feature Selection

The initial step in feature selection involved the determination of a loss function to identify the most suitable features. This function was established by evaluating the performance of decision tree models using 42 initial features within the initial phase of model evaluation. Performance metrics of ML algorithms during this stage are detailed in Table 2.

The findings indicated that the Extra Trees Classifier model, exhibiting the highest AUC, outperformed other models. Consequently, the loss function was defined based on the accuracy of this particular model. Meta-heuristic algorithms were employed for feature selection based on the Loss function. These algorithms are designed to minimize the loss function and systematically search for a binary list with the lowest value. The results of using a list of the most important meta-heuristic algorithms are shown in supplementary information (Table 1). Furthermore, the evaluation of four meta-heuristic algorithms based on their performance revealed that the bat algorithm outperformed the others by achieving the highest score of 89.5% (0.895) in iteration 128 (Figure 3). The population data is refreshed, and the primary iteration concludes, repeating until the specified termination criterion is satisfied. In this particular scenario, the stopping criterion was defined as reaching a total of 200 iterations. The bat algorithm identified a total of 20 features for selection. The feature selection process is illustrated in supplementary information (Figure 1).

Table 1. Patient's characteristics

Variables	Scale	Class dose diazepam				
			Patients who received >10 mg diazepam (232)	Patients who received >10 mg diazepam (383)	Total (615)	p-value
Gender (n, %)	Nominal	Male	224 (38%)	365 (62%)	589	0.455
		Female	8 (31%)	18 (69%)	26	
Age (median, IQR)	Interval	------	34 (18)	37 (16)		0.149
Purpose of naltrexone use	Nominal	Opioid addiction treatment	3 (16%)	15 (84%)	18	0.019
		Suicide	52 (48%)	57 (52%)	109	
		Accident	5 (23%)	16 (77%)	21	
		Unknown	172 (36%)	295 (62%)	476	
Temperature	Interval	----	37 (0.1)	36.9 (0.1)	----	0.14
Systolic blood pressure	Interval	----	120 (20)	120 (15)	----	0.742
Diastolic blood pressure	Interval	----	75 (10)	75 (10)	----	0.902
$SpaO_2$	Interval	----	97 (1)	97 (0)	----	0.582
Vomiting	Nominal	Yes	35 (30%)	80 (70%)	115	0.07
		No	197 (39%)	303 (61%)	500	
Nausea	Nominal	Yes	37 (34%)	70 (66%)	107	0.46
		No	195 (38%)	313 (62%)	508	
Diarrhea	Nominal	Yes	19 (37%)	32 (63%)	51	0.94
		No	213 (38%)	351 (62%)	564	
Yawning	Nominal	Yes	5 (38%)	8 (625)	13	0.95
		No	277 (46%)	375 (54%)	602	
Lacrimation	Nominal	Yes	2 (18%)	9 (82%)	11	0.17
		No	230 (38%)	374 (62%)	604	
Respitatory	Nominal	Yes	3 (27%)	8 (73%)	11	0.47
		No	229 (38%)	375 (62%)	604	
ECG_Rate	Interval	---	80 (15)	80 (8)	---	0.95
Rhythm	Nominal	Normal sinus	182 (36%)	317 (64%)	499	0.324
		Not-sinus-AF	23 (44%)	29 (56%)	52	
		AF	24 (44%)	31 (56%)	55	
		Sinus tachycardia	1 (25%)	3 (75%)	4	
		Sinus bradycardia	2 (40%)	3 (60%)	5	
Axis deviation	Nominal	Normal	213 (38%)	348 (62%)	561	0.445
		Right axis deviation	12 (43%)	16 (57%)	28	
		Left axis deviation	7 (27%)	19 (73%)	26	
ST elevation	Nominal	Yes	2 (33%)	4 (67%)	6	0.82
		No	230 (38%)	379 (62%)	609	
T invert	Nominal	Yes	5 (28%)	13 (72%)	18	0.377
		No	227 (46%)	370 (54%)	597	
T flat	Nominal	Yes	5 (55%)	4 (45%)	9	0.266
		No	227 (37%)	379 (63%)	606	
VBG_PH	Interval	----	7.41 (0)	7.41 (0)	----	0.34
VBG_Pco_2	Interval	----	40 (0)	40 (0)	----	0.36
VBG_BE	Interval	----	0.7 (0)	0.7 (0)	----	0.05
p	Interval	----	80 (13)	80 (12)	----	0.54
K	Interval	----	4.1 (0.4)	4.1 (0.5)	----	0.722

Table 1. Continued

Variables	Scale	Class dose diazepam				
			Patients who received >10 mg diazepam (232)	Patients who received <10 mg diazepam (383)	Total (615)	p-value
BS	Interval	----	107 (24)	108 (23)	----	0.44
VBG_Po_2	Interval	----	38.4 (0)	38.4 (0)	----	0.629
VBG_Hco_3	Interval	----	24.5 (0)	24.5 (0)	----	0.145
Hgb	Interval	----	13.9 (1.45)	13.9 (1.3)	----	0.756
Na	Interval	----	141 (4)	141 (4)	----	0.759
Cr	Interval	----	1.04 (0.2)	1.04 (0.2)	----	0.691
AIT	Interval	----	38 (0)	38 (0)	----	0.004
BUN	Interval	----	30 (13)	30 (10)	----	0.311
AST	Interval	----	34 (0)	34 (0)	----	0.075
ALK	Interval	----	214 (0)	214 (0)	----	0.654
CK	Interval	----	666 (0)	666 (0)	----	0.064
LDH	Interval	----	662 (0)	662 (0)	----	0.192
Seizure	Nominal	Yes	2 (13%)	13 (87%)	15	0.04
		No	230 (38%)	370 (62%)	600	
Recent opioid use, <1 week, (n, %)	Nominal	Yes	148 (91%)	13 (9%)	161	0.000
		No	84 (18%)	370 (82%)	454	
Time elapsed before hospital admission	Interval	----	2 (4)	3 (3)		0.000
Naltrexone intake quantity (mg)	Interval	----	50 (50)	50 (0)		0.000
Level of consciousness	Ordinal	Conscious	183 (38%)	297 (62%)	480	0.86
		Grade 1	40 (37%)	69 (73%)	109	
		Grade 2	7 (37%)	12 (73%)	19	
		Grade 3	2 (40%)	3 (60%)	5	
		Grade 4	0	2 (100%)	2	
RASS	Ordinal	---	2 (4)	0 (2)	---	0.00

IQR: Interquartile range, ECG: Electrocardiogram, ST: Segment, AF: Atrial fibrillation, K: Potassium, BS: Blood sugar, Hgb: Hemoglobin, Na: Sodium, Cr: Creatinine, AIT: Autoimmune thyroiditis, BUN: Blood urea nitrogen, AST: Aspartate aminotransferase, ALK: Alkaline phosphatase, CK: Creatine kinase, LDH: Lactate dehydrogenase, RASS: Richmond agitation-sedation scale

Table 2. Ten-fold cross-validation for classifiers performance on selected predictors in classification decision tree family (42 features) in train datasets

Model		Accuracy	AUC	Sensitivity	Precision	F1
Light gradient boosting machine	Train	0.9186	0.9290	0.9186	0.9225	0.9169
	Test	0.8919	0.9210	0.8919	0.8919	0.8919
Random Forest Classifier	Train	0.9047	0.9241	0.9047	0.9079	0.9031
	Test	0.8757	0.9201	0.8757	0.8761	0.8758
Gradient Boosting Classifier	Train	0.9023	0.9241	0.9023	0.9054	0.9006
	Test	0.8865	0.9098	0.8865	0.8862	0.8863
Extreme Gradient Boosting	Train	0.9000	0.9252	0.9000	0.9030	0.8983
	Test	0.8919	0.9241	0.8919	0.8919	0.8919
Extra Trees Classifier	Train	0.8977	0.9226	0.8977	0.9017	0.8955
	Test	0.8703	**0.9306***	0.8703	0.8695	0.8695

AUC: Area under the curve, *: Maximum value

Model Evaluation (Stage 2)

The outcomes of ten-fold cross-validation for the performance of classifiers on a set of 20 features in the classification decision tree family are presented in Table 3. The analysis indicates that Extreme Gradient Boosting (XGBoost) achieved the highest levels of accuracy (0.8811), sensitivity (0.8811), precision (0.8819), and F1-score (0.8814) on the test dataset. Furthermore, the Extra Trees Classifier exhibited the highest AUC compared to other decision tree classifiers (AUC: 0.9343). Based on the findings in Table 3, the Extra Trees Classifier model emerged as the most suitable choice for further investigation into feature importance for weighting neural network performance through model tuning. The observed discrepancy between values of model's performance metrics was due to feature selection. Models trained on a reduced set of 20 selected features demonstrated superior performance compared to models using the original 42 features, highlighting the impact of feature selection on predictive accuracy.

Feature Weight Calculation Using Decision Trees

To determine the importance coefficient of 20 chosen features, researchers utilized the assistance of the Extra Trees Classifier, selected based on the optimal AUC metric obtained during model evaluation. The top 20 features and their respective importance values are detailed in Figure 4. The figure illustrates the ranking of feature importance in a descending order on the y-axis, with the x-axis representing the corresponding importance values. The researchers identified and ranked the most significant features in the following sequence: recent opioid use, RASS score, naltrexone intake quantity, heart rate, systolic blood pressure, time elapsed before hospital admission, sodium level, creatinine level, level of consciousness, VBG-HCO_3, purpose of naltrexone use, nausea, lactate dehydrogenase, AXIS, $SpaO_2$, seizure occurrence, gender, T-flat wave, respiratory rate, and lacrimation (Figure 4).

Neural Network Results

The outcomes of incorporating weighted features into a MLP neural network were detailed in Table 4. The MLP model demonstrated notable performance metrics, including an AUC of 0.94, accuracy of 0.91, precision of 0.92, sensitivity of 0.9, and an F1-score of 0.91. Additionally, Figures 5 to 7 visually depict the enhanced effectiveness of the weighted MLP neural network when compared with ML techniques like Extreme Gradient Boosting and Extra Trees Classifier. The values reported in the Tables 3 and 4 report the average values of evaluation metrics calculated for the two study groups. The minor discrepancies between the values shown in Figures 6 and 7 and those in Tables 3 and 4 is because the figures report numbers rounded to two decimal places.

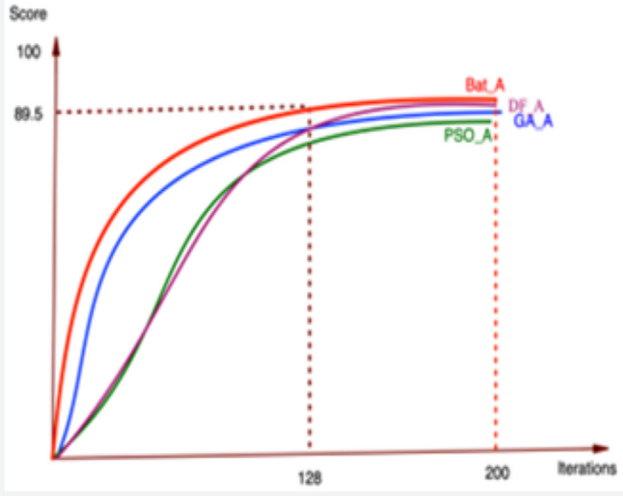

Figure 3. Performance of four meta-heuristic algorithms

Table 3. Ten-fold cross-validation for classifiers performance on selected predictors in classification decision tree family (20 features) in train datasets

Model		Accuracy	AUC	Sensitivity	Precision	F1
Gradient Boosting Classifier	Train	0.9256	-	0.9256	0.9303	0.9237
	Test	0.8703	0.9227	0.8703	0.8697	0.8699
Extreme Gradient Boosting	Train	0.9209	-	0.9209	0.9254	0.9201
	Test	**0.8811***	0.9175	**0.8811***	**0.8819***	**0.8814***
Light Gradient Boosting Machine	Train	0.9140	-	0.9140	0.9148	0.9129
	Test	0.8703	0.9252	0.8703	0.8703	0.8703
Random Forest Classifier	Train	0.9116	-	0.9116	0.9128	0.9107
	Test	0.8703	0.9302	0.8703	0.8697	0.8699
Extra Trees Classifier	Train	0.9070	-	0.9070	0.9093	0.9056
	Test	0.8649	**0.9343***	0.8649	0.8645	0.8649

*: Maximum value, AUC: Area under the curve

DISCUSSION

This research employed metaheuristic-based algorithms for the purpose of feature selection. Metaheuristic algorithms are a type of optimization methods used to tackle complex optimization problems that traditional approaches may have difficulty solving effectively. These algorithms are known for their adaptability in various optimization fields, including engineering, logistics, finance, and artificial intelligence[18]. Metaheuristic algorithms excel at identifying the most suitable subset of features in a dataset while maintaining model accuracy. Considering their effectiveness, this research focuses on leveraging metaheuristic algorithms to address feature selection complexities[19]. These algorithms efficiently solve complex optimization problems. Their significance is particularly notable in complex medical scenarios, such as diagnosis and treatment, especially in domains like drug dosage determination where predictive variables may be scarce. The utilization of metaheuristic algorithms to address these optimization challenges has emerged as a promising approach for handling NP-hard problems.

Modern challenges require quick solutions, making classical approaches inadequate. This has led to the rise of meta-heuristic algorithms, which explore spaces efficiently using a single fitness function, often with swarm intelligence. These algorithms can be population-based, like Genetic Algorithm, or path-based, like bat algorithm, which excels in complex biomedical scenarios. In this study, the bat algorithm outperformed other methods in feature selection, showing promise for solving challenging optimization problems in various fields[20,21].

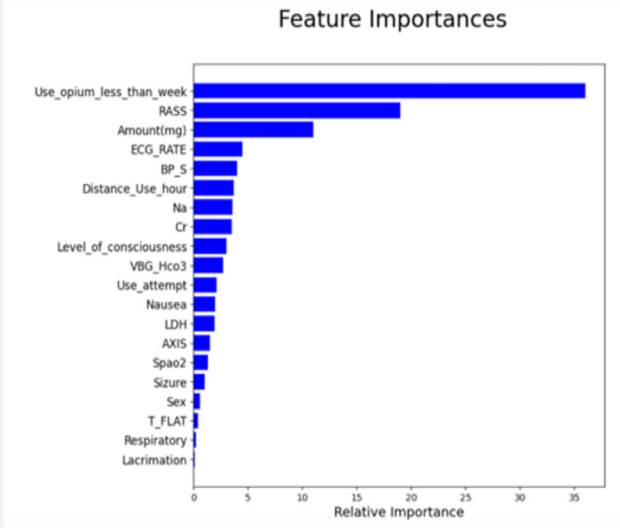

Figure 4. Feature importance

ECG: Electrocardiogram, Na: Sodium, Cr: Creatinine, LDH: Lactate dehydrogenase, RASS: Richmond Agitation-sedation scale, VBG: Venous blood gases

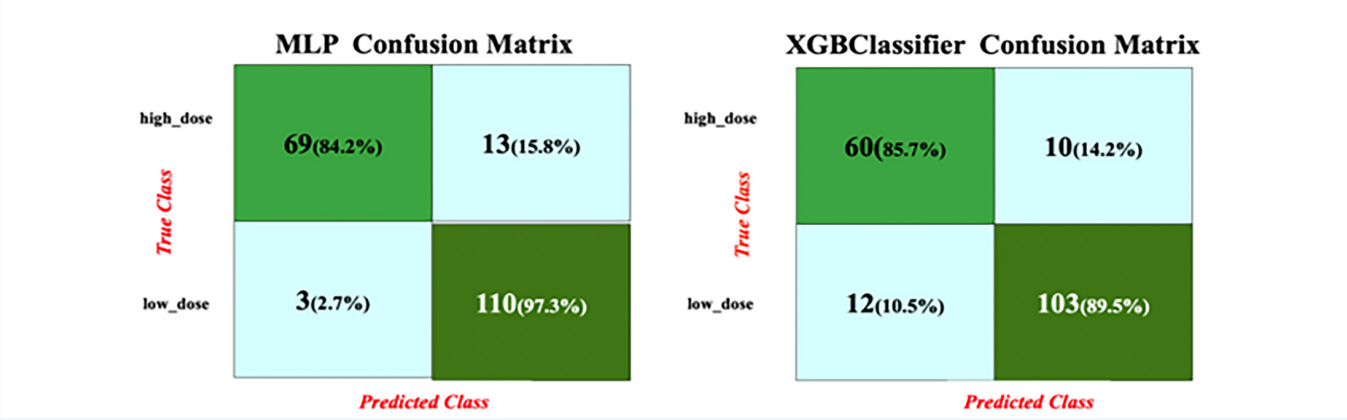

Figure 5. Confusion matrix comparison between the optimal machine learning classifier model, XGBoost, and the weighted MLP neural network model

MLP: Multilayer perceptron

Table 4. The outcomes of incorporating weighted features into a multilayer perceptron neural network

Model	Dataset	Accuracy	AUC	Sensitivity	Precision	F1
MLP	Train	**0.93**	-----	**0.92**	**0.94**	**0.93**
	Test	**0.91**	**0.94**	**0.90**	**0.92**	**0.91**

AUC: Area under the curve, MLP: Multilayer perceptron

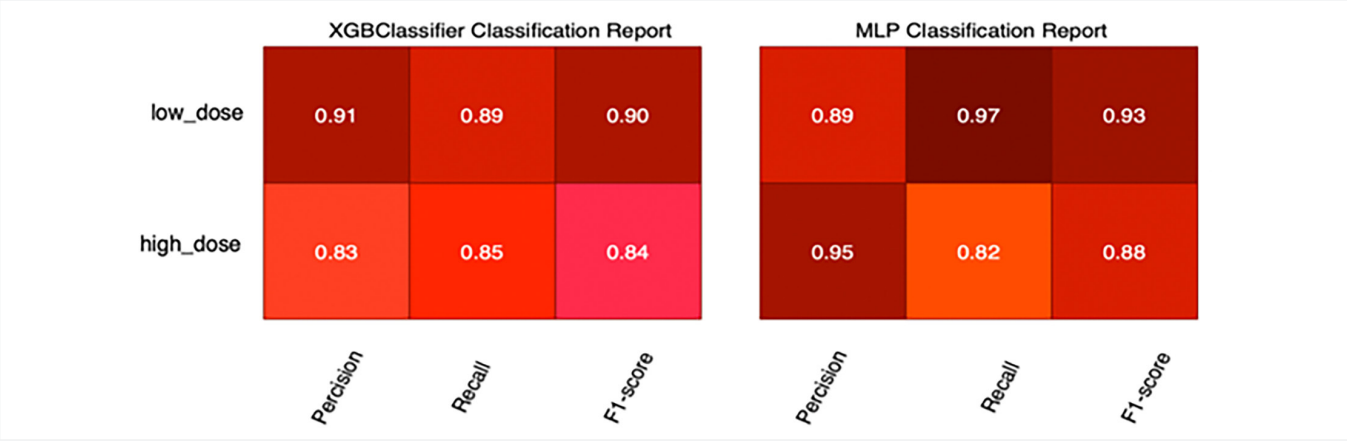

Figure 6. Comparison of the precision, sensitivity, and F1-score metrics in the optimal machine learning model (XGBoost) and weighted MLP neural network model

MLP: Multilayer perceptron

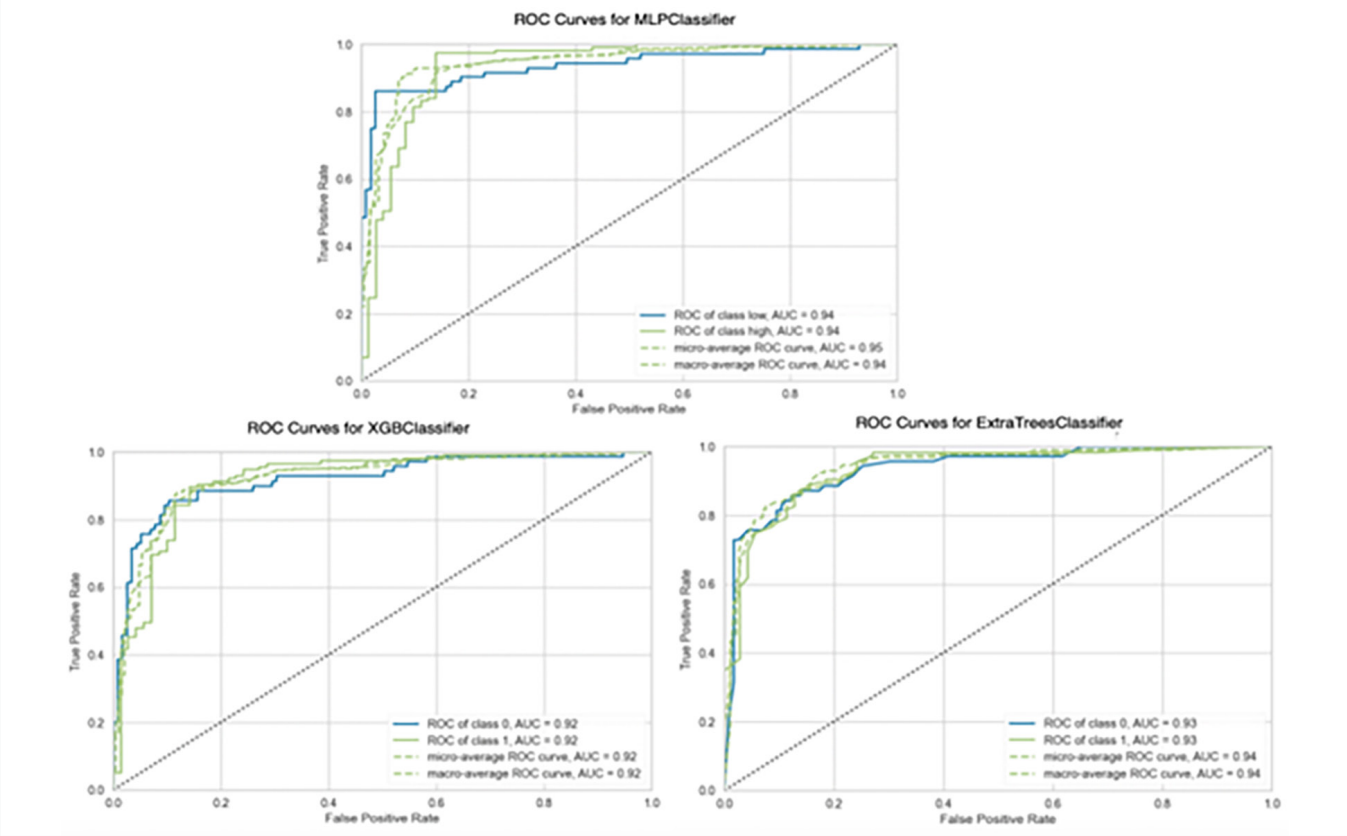

Figure 7. Comparison of the ROC curve between two machine learning models (XGBoost and Extra Trees classifier) and weighted MLP neural network model

MLP: Multilayer perceptron, AUC: Area under the curve

Among developed ML models, the Extra Trees Classifier yield the strongest performance based on its AUC score. The Extra Trees algorithm excels at selecting informative features via gradient-based techniques, enhancing predictive modeling outcomes. For example, Mathpal et al.[22] reported how Extra Trees played a central role in identifying compounds targeting mutant PBP4 in *Staphylococcus aureus*, achieving an 81% predictive accuracy. This highlights the model's effectiveness in not only making predictions but also elucidating which features contribute most significantly, thus aiding in clinical decision-making and research explorations. Numerous studies demonstrate that predictions using Extra Trees Classifiers have been successfully applied across various critical healthcare settings, from early detection of diabetes[23] to the prediction of abdominal aortic aneurysms[24]. The ability to predict outcomes reliably emphasizes the model's critical role in improving patient management and therapeutic strategies. A study compared decision tree classifiers with a weighted MLP neural network, finding that the MLP network showed superior performance in AUC and other evaluation criteria. Despite typically requiring large training samples, MLP models can be accurate with limited data by restricting input parameters and utilizing feature weight calculation from the highest AUC classifier. The ROC curve, unaffected by imbalanced data, makes AUC a preferred metric for evaluating ML models[25,26].

Recent opioid use was found to be the primary factor influencing the necessary dosage of diazepam to manage agitation triggered by naltrexone. The absorption of naltrexone through oral intake occurs swiftly, reaching peak levels in the bloodstream within 3 hours. It is advised that individuals be free of opioids for 7 to 10 days before receiving naltrexone. Consequently, patients who have used opioids within a week prior to naltrexone administration are at a higher risk of encountering acute opioid receptor blockade and intense opioid withdrawal symptoms, necessitating increased diazepam dosages[9]. A higher RASS score was found to be a strong indicator of a high dosage of diazepam, consistent with prior research that has demonstrated a relationship between RASS score and the amount of sedative and analgesic drugs administered[27,28]. A higher RASS score may signal a need for >10 mg diazepam to achieve sedation. However, the exact RASS cut-off value needs to be determined in future studies. Increased consumption of naltrexone was linked to an increased need for diazepam to alleviate agitation, likely attributable to its competitive antagonistic properties and the dose-dependent manner in which it blocks opioid receptors[29]. Although, in the study conducted by Sabzghabaee et al.[11] it has been reported that diazepam cannot effectively reduce agitation until 120 minutes post-administration, and its mean onset of action is lower than midazolam. In a systematic review by Kunzler et al.[12], reported diazepam dosing regimens varied by treatment scenario: 5-10 mg for oral naltrexone, 10 mg for extended-release injectable naltrexone (Vivitrol®), up to 30 mg for naltrexone implants or nalmefene (18 mg), and 10-40 mg for high-dose naltrexone (50 mg).

Several factors such as elevated pulse rate, increased systolic blood pressure, heightened level of consciousness, presence of nausea, reduced lacrimation, and decreased need for mechanical intubation are more prevalent in severely agitated patients, potentially necessitating higher doses of diazepam for resolution. Gender was identified as a feature; further research is needed to explore the potential relationship between gender, agitation intensity, and diazepam dosage requirements. It is important to note that suicide is a complex issue with multiple contributing factors. Additional predictors warranting further investigation for their potential association with increased diazepam dosages in these patients include ECG abnormalities (such as T-wave flattening and axis deviation), levels of sodium and creatinine, arterial oxygen saturation, and bicarbonate levels in VBG analysis.

Study Limitations

There are several limitations that need to be acknowledged. Primarily, the study was constrained by the challenges associated with data collection from multiple medical facilities in Iran, resulting in the utilization of data from a single hospital. As regional differences in opioid use patterns may affect generalizability, future investigations should consider expanding the sample size or incorporating data from multi-center datasets. Furthermore, the study was restricted to a limited selection of five ML models. To gain a more comprehensive understanding, it is advisable for subsequent research to explore a broader array of models. Moreover, the method to handle the missing data may introduce bias. For instance, mean imputation, may underestimate variability for variables with non-random missingness and stochastic regression imputation relies on correctly specified regression models. Future work could explore advanced methods (e.g., Bayesian imputation) for complex missingness patterns. Lastly, the retrospective nature of the dataset may introduce biases and uncertainties stemming from missing data.

CONCLUSION

Our research findings indicate that the utilization of a weighted MLP neural network can be an effective tool in developing a prediction model for the necessity of higher doses of diazepam for patients experiencing naltrexone-induced agitation. Particularly noteworthy is the enhanced predictive capability achieved when employing meta-heuristic techniques for feature selection and subsequently weighting these selected features using a classifier with the highest AUC value. In conclusion, our study emphasizes the value

of employing a weighted MLP neural network to improve predictive accuracy and facilitate the tailored management of patients experiencing naltrexone-induced agitation and could reduce trial-and-error dosing, improving patient safety and staff efficiency.

Ethics

Ethics Committee Approval: The study received approval from the Ethics Committee of Shahid Beheshti University of Medical Sciences (decision no: IR.SBMU.RETECH.REC.1402.626, date: 01.07.2024).

Informed Consent: This study is a retrospective cross-sectional analysis of medical records of patients experiencing naltrexone-induced withdrawal symptoms at Loghman Hakim Hospital from April 2002 to March 2016.

Footnotes

Authorship Contributions

Concept: S.A.M., S.M.H., S.S., Design: S.A.M., Data Collection or Processing: S.A.M., Analysis or Interpretation: S.A.M., Literature Search: S.A.M., S.S., S.F., O.M., B.M., P.E.T.E., M.R., Writing: S.A.M., S.M.H.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Goel A, Kapoor B, Wu M, Iyayi M, Englesakis M, Kohan L, et al. Perioperative naltrexone management: a scoping review by the perioperative pain and addiction interdisciplinary Network. Anesthesiology. 2024;141:388-99.
2. Zangiabadian M, Golmohammadi S, Nejadghaderi SA, Zahmatkesh MM, Nasiri MJ, Sadeghian M. The effects of naltrexone on retention in treatment and being opioid-free in opioid-dependent people: a systematic review and meta-analysis. Front Psychiatry. 2022;13:1003257.
3. Gibson A, Degenhardt L. Mortality related to naltrexone in the treatment of opioid dependence: a comparative analysis NDARC Technical Report No. 229.
4. Rahimi M, Kargar A, Hazegh Fetratjoo D, Hosseini SM, Mahdavinejad A, Shadnia S. Demographic and clinical characteristics of 907 cases with naltrexone intoxication; a 14-year cross-sectional study. Arch Acad Emerg Med. 2022;10:e34.
5. Saadabadi DSA. Naltrexone StatPearls Treasure Island (FL): StatPearls Publishing. Available from: https://www.ncbi.nlm.nih.gov/books/NBK534811/.
6. Erlich Y. Modulators, mediators, and specifiers in brain function: interactions of neuropeptides, cyclic nucleotides, and phosphoproteins in mechanisms underlying neuronal activity, behavior, and neuropsychiatric disorders. Springer Science & Business Media. 2012.
7. Volavka J, Mallya A, Bauman J, Pevnick J, Cho D, Reker D, et al. Hormonal and other effects of naltrexone in normal men. Adv Exp Med Biol. 1979;116:291-305.
8. Darke S, Farrell M, Duflou J, Larance B, Lappin J. Circumstances of death of opioid users being treated with naltrexone. Addiction. 2019;114:2000-7.
9. Hassanian-Moghaddam H, Afzali S, Pooya A. Withdrawal syndrome caused by naltrexone in opioid abusers. Hum Exp Toxicol. 2014;33:561-7.
10. Aggarwal K, Ranjan R, Mallik S, Salian HH, Shekhar S. Delirium secondary to inadvertent administration of naltrexone in patient with opioid dependence syndrome: a case report. East Asian Arch Psychiatry. 2024;34:87-8.
11. Sabzghabaee AM, Eizadi-Mood N, Gheshlaghi F, Javani A, Shirani S, Aghaabdollahian S. Role of benzodiazepines in the management of agitation due to inappropriate use of naltrexone. Iran J Nurs Midwifery Res. 2012;17:365-9.
12. Kunzler NM, Wightman RS, Nelson LS. Opioid withdrawal precipitated by long-acting antagonists. J Emerg Med. 2020;58:245-53.
13. Rund DA, Ewing JD, Mitzel K, Votolato N. The use of intramuscular benzodiazepines and antipsychotic agents in the treatment of acute agitation or violence in the emergency department. J Emerg Med. 2006;31:317-24.
14. Abdel-Basset M, Abdel-Fatah L, Sangaiah AK. Metaheuristic algorithms: a comprehensive review. Computational intelligence for multimedia big data on the cloud with engineering applications. 2018:185-231.
15. Singh J, Sandhu JK, Kumar Y. Metaheuristic-based hyperparameter optimization for multi-disease detection and diagnosis in machine learning. Service Oriented Computing and Applications. 2024;18:163-82.
16. Parizi MK. Binary modified cat and mouse based optimizer for medical feature selection: a COVID-19 case study. 2024.
17. Mohtarami SA, Mostafazadeh B, Shadnia S, Rahimi M, Evini PET, Ramezani M, et al. Prediction of naloxone dose in opioids toxicity based on machine learning techniques (artificial intelligence). Daru. 2024;32:495-513.
18. Watgule U. Metaheuristics Algorithms 2023. Available from: https://medium.com/@ujwalwatgule/metaheuristics-algorithms-889d2b0f9652.
19. Akinola OO, Ezugwu AE, Agushaka JO, Zitar RA, Abualigah L. Multiclass feature selection with metaheuristic optimization algorithms: a review. Neural Comput Appl. 2022;34:19751-90.
20. Sadeghian Z, Akbari E, Nematzadeh H, Motameni H. A review of feature selection methods based on meta-heuristic algorithms. Journal of Experimental & Theoretical Artificial Intelligence. 2023:1-51.
21. Kar AK. Bio inspired computing - a review of algorithms and scope of applications. Expert Systems with Applications. 2016;59:20-32.
22. Mathpal S, Joshi T, Priyamvada P, Ramaiah S, Anbarasu A. Machine learning and cheminformatics-based Identification of lichen-derived compounds targeting mutant $PBP4^{R200L}$ in Staphylococcus aureus. Mol Divers. 2025;29:3345-70.
23. Arya M, Sastry G H, Motwani A, Kumar S, Zaguia A. A novel extra tree ensemble optimized DL framework (ETEODL) for early detection of diabetes. Front Public Health. 2022;9:797877.
24. Forneris A, Beddoes R, Benovoy M, Faris P, Moore RD, Di Martino ES. AI-powered assessment of biomarkers for growth prediction of abdominal aortic aneurysms. JVS Vasc Sci. 2023;4:100119.
25. Fujita T, Sato A, Narita A, Sone T, Iokawa K, Tsuchiya K, et al. Use of a multilayer perceptron to create a prediction model for dressing independence in a small sample at a single facility. J Phys Ther Sci. 2019;31:69-74.
26. Siciarz P, Alfaifi S, Uytven EV, Rathod S, Koul R, McCurdy B. Machine learning for dose-volume histogram based clinical decision-making support system in radiation therapy plans for brain tumors. Clin Transl Radiat Oncol. 2021;31:50-7.
27. Lee JA, Duby JJ, Cocanour CS. Effect of early and focused benzodiazepine therapy on length of stay in severe alcohol withdrawal syndrome. Clin Toxicol (Phila). 2019;57:624-7.
28. Ely EW, Truman B, Shintani A, Thomason JW, Wheeler AP, Gordon S, et al. Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation-Sedation Scale (RASS). JAMA. 2003;289:2983-91.
29. Sirohi S, Dighe SV, Madia PA, Yoburn BC. The relative potency of inverse opioid agonists and a neutral opioid antagonist in precipitated withdrawal and antagonism of analgesia and toxicity. J Pharmacol Exp Ther. 2009;330:513-9.