Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 29s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Medicine Overdose Prediction Using Machine Learning

Mr. S.Sedhumadhavan[1], K. Gomathi[2], M. Imayajothy[3], A.P. Kamaleishuvari[4]

[1] Assistant Professor (Stage-III), Department of Biomedical Engineering Rajiv Gandhi College of Engineering and Technology, Puducherry, India.

Email ID: sedhumadhavan@rgcet.edu.in

[2] Department of Biomedical Engineering, Rajiv Gandhi College of Engineering and Technology, Puducherry, India.

Email ID: gomathikarthika18@gmail.com

[3] Department of Biomedical Engineering, Rajiv Gandhi College of Engineering and Technology, Puducherry, India.

Email ID: imayajothy@gmail.com

[4] Department of Biomedical Engineering, Rajiv Gandhi College of Engineering and Technology, Puducherry, India.

Email ID: kamaleishuvari2004@gmail.com

Cite this paper as: Mr. S.Sedhumadhavan, K. Gomathi, M. Imayajothy, A.P. Kamaleishuvari, (2025) Medicine Overdose Prediction Using Machine Learning. *Journal of Neonatal Surgery*, 14 (29s), 35-40.

ABSTRACT

The opioid addiction crisis in the United States has gained national attention due to the alarming rise in opioid overdose (OD)-related deaths. Addressing this epidemic has become a top priority for governments and healthcare providers, requiring critical insights into the risk factors associated with opioid overdose. In this paper, we present our work on developing machine learning-based prediction models to assess the likelihood of opioid overdose using patients' electronic health records (EHR).

We conducted two studies utilizing New York State claims data (SPARCS) with 440,000 patients and Cerner's Health Facts database with 110,000 patients. Our experiments demonstrated that EHR-based prediction models achieved the highest recall using the random forest method (precision: 95.3%, recall: 85.7%, F1 score: 90.3%), while deep learning yielded the highest pre- cision (precision: 99.2%, recall: 77.8%, F1 score: 87.2%). Addi- tionally, we identified clinical events as key features contributing to the accuracy of these predictions.

Keywords: *Medicine Overdose, Machine Learn- ing, Predictive Analytics, Healthcare AI, Risk Assessment.*

1. INTRODUCTION

The United States faces an opioid epidemic, with misuse of prescribed pain relievers and illegal opioids like heroin and fentanyl leading to a surge in overdose-related deaths. In 2017, drug overdoses caused 70,237 deaths, with 67.8 percent involving opioids. Improving clinical practices, such as CDC guidelines and pain management standards, can help mitigate risks. Identify applicable funding agency here. If none, delete this. Clinical Decision Support Systems (CDSS) leverage Electronic Health Records (EHR) to enhance decision-making and identify individuals at risk of opioid toxicity The widespread adoption of EHR, supported by initiatives like HITECH, has enabled large-scale data analysis. Government and commercial EHR databases, such as SPARCS and Cerner's Healthcare Data Insights, provide valuable resources for predictive modeling. Machine learning, including deep learning techniques, has been increasingly applied to EHR-based predictive modeling. Studies have used deep neural networks and recurrent neural networks (RNN) for disease prediction and opioid use classification. Our research developed multiple models using claims and EHR data to predict opioid poisoning risks. By analyzing demographic information, medical history, diagnoses, and clinical events, we identified key predictive features. Our models achieved high accuracy, with Random Forest yielding a precision of 95.3 percent and recall of 85.7percent, while Neural Networks achieved a precision of 99.2 percent and an UC of 95.4 percent. These findings demonstrate the potential of machine learning in opioid risk prediction, highlighting the importance of clinical event data in enhancing predictive accuracy.

I. RELATED WORK

A. *Prediction Results Analysis*

our experiments, we utilized SPARCS and Health Facts databases., splitting 90% both positive and negative cases for training and 10% for testing. For SPARCS, the training set included 360,000 negative and 36,000 positive cases, while the test set had 40,000 negative and 4,000 positive cases. For Health Facts, the training set had 90,000 negative and 9,000 positive cases, with 10,000 negative and 1,000 positive cases in the test set. To assess model performance, we measured accuracy, precision, recall, F1-score, and AUC. Given the dataset im-balance, recall is crucial, and F1-score provides an aggregated performance measure. We compared traditional and deep learning models, with results detailed in Table 2, where the best-performing values are highlighted. Results demon- strated our models are capable of classifying opioid poisoning very well. Our models perform effectively, achieving an F1 score of 79.27% for SPARCS using deep learning and 90.30% for Health Facts with the random forest method. Deep learning yields the highest precision, reaching 97.07% for SPARCS and 99.23% for Health Facts, highlighting the models' ability to accurately identify opioid-poisoned patients. In terms of

recall, the models achieve 71.68% for SPARCS and 85.70% for Health Facts, these recall values represent the model's ability to identify all opioid-poisoned patients in the dataset. The highest AUC scores achieved are 94.94% for SPARCS and 95.41% for Health Facts. In the SPARCS dataset, which contains fewer features, deep learning produces the best F1 score while achieving an AUC score comparable to that of random forest. In contrast, for Health Facts, which includes a more extensive set of features, random forest delivers the highest F1 score, whereas deep learning attains the best AUC score.

TABLE I: Experiment Result of different methods on SPARCS and Health Facts datasets

Dataset	Model	Acc.	Prec.	Rec.	F1	AUC
SPARCS	Random Forest	96.03%	88.03%	64.25%	74.61%	94.94%
	Decision Tree	95.33%	75.37%	71.68%	73.17%	94.62%
	Logistic Reg.	96.44%	96.95%	67.85%	76.22%	81.33%
	Deep Learning	**96.82%**	**97.07%**	**72.02%**	**79.27%**	**94.92%**
Health Facts	Random Forest	98.69%	95.34%	85.78%	90.30%	95.11%
	Decision Tree	97.83%	87.47%	83.10%	85.23%	93.83%
	Logistic Reg.	95.73%	96.17%	55.20%	70.14%	74.32%
	Deep Learning	**97.93%**	**99.43%**	**77.80%**	**87.22%**	**95.41%**

TABLE II: Top 20 Important Features for Prediction on SPOCS Dataset

Category	Description	Rank	Category	Description	Rank
Diagnosis	Sedative dep.	1	Diagnosis	Altered mental status	11
Diagnosis	Viral hep. C	2	Diagnosis	Drug abuse, unsp.	12
Diagnosis	Consciousness alt.	3	Diagnosis	Chronic pain	13
Diagnosis	Cannabis abuse	4	Diagnosis	Respiratory failure	14
Diagnosis	Anxiety state	5	Diagnosis	Lumbago	15
Diagnosis	Drug dep.	6	Diagnosis	CNS stimulant poison	16
Diagnosis	Sedative abuse	7	Diagnosis	Depressive dis-order	17
Diagnosis	Alcohol toxic-ity	8	Diagnosis	Benzodiazepine poison	18
Diagnosis	Chronic pain	9	Diagnosis	Tobacco use disorder	19
Diagnosis	Pedestrian injury	10	Procedure	Tissue excision (deep)	20

B. *Feature Analysis*

Understanding the significance of various features in the models is crucial for helping researchers and clinicians ex- plore potential causes or disease progression patterns. Using our best-performing traditional method, random forest, we identified the most important features for both datasets, as presented in Tables 3 and 4. These tables also include the ranking and feature categories. Random forest determines feature importance based on Gini importance, which measures the total reduction in node Gini impurity index, weighted by the probability of reaching that node, and averaged across all decision trees in the ensemble. The Gini impurity index is defined as:

$G = i = 1ncpi(1pi)eq$ (1)

Random Forest is a collection of decision trees, where each tree classifies input samples as positive or negative. The final decision is determined by combining results from all trees. Each decision tree node selects a feature to determine the path, affecting class distribution and Gini impurity index. The Gini importance is then derived based on these variations.

$I = GparentGsplit1Gsplit2eq$ (2)

The Gini impurity index measures node purity in decision trees, with Gini importance representing a feature's contribu- tion across all trees. For the New York State SPARCS dataset, diagnosis codes are the most significant predictive features, with sedative, hypnotic, or anxiolytic dependence ranking as the top feature (ICD-9:304.11, ICD-10: F13.20). Other key features include Hepatitis C, altered consciousness, cannabis

abuse, and chronic pain conditions. The Health Facts database, with richer clinical data, high- lights heart rate, respiratory rate, temperature, pain scale, fall history, and BMI as top predictors, as opioid overdose affects vital signs. Table 5 summarizes feature importance, with diagnosis dominating SPARCS predictions, followed by procedures and demograph- ics, while clinical events lead in Health Facts, followed by diagnosis, medications, and procedures.

TABLE III: Importance for Each Feature Category

SPARCS			Health Facts				
Diag.	**Proc.**	**Demo.**	**Diag.**	**Proc.**	**Meds**	**Clin.**	**Demo.**
78.1%	15.3%	6.6%	25.1%	6.2%	13.8%	50.1%	4.8%

C. Experiment on Predicting Outcomes Using Only Primary Data

To evaluate performance using only primary data, we re- moved diagnosis codes from the Health Facts dataset and rebuilt the models. Table 6 compares results before and after removal. Despite the exclusion, prediction performance remained strong, with Random Forest achieving a slightly higher AUC was achieved without including diagnosis codes.

TABLE IV: Performance of Models Before and After Removing Diagnostic Features on Health Facts

Methods	Diag. Codes	Accuracy	Precision	Recall	F1	AUC
Random Forest	Yes	98.69%	95.34%	85.70%	90.30%	95.11%
	No	97.41%	96.25%	74.10%	85.36%	95.29%
Decision Tree	Yes	97.38%	87.47%	73.10%	82.29%	94.53%
	No	96.98%	93.15%	72.10%	83.29%	94.58%
Logistic Regression	Yes	95.73%	96.66%	55.80%	70.10%	95.69%
	No	95.73%	96.63%	55.80%	70.10%	95.69%
Deep Learning	Yes	97.35%	98.36%	72.00%	83.14%	94.97%
	No	97.35%	98.36%	72.00%	83.14%	94.97%

A. Data Sources

2. METHODOLOGY

We utilized SPARCS and Cerner's Health Facts as a data source, while SPARCS inpatient data includes hospital dis- charge records from New York State healthcare facilities, sup- porting statewide healthcare analysis. Health Facts inpatient data comprises Anonymized EHR records from over 600 U.S. hospitals, offering a comprehensive dataset with diagnoses, procedures, demographics, medications, vital signs, lab results, and other clinical observations.

B. Diagnosis Codes for Data Retrieval

Patients were categorized into opioid overdose and non- opioid groups based on opioid-related diagnosis codes in their records. ICD-9 and ICD-10 codes were used to identify poi- soning from opiates, heroin, methadone, and other narcotics. Since future opioid overdose risk is unknown, a patient's last visit status was used as the prediction target, while prior visit data served as features for model training.

C. Research Datasets

We created two separate study datasets: SPARCS (2005–2016) and Health Facts (2000–2017). In SPARCS, we selected 40,000 opioid-poisoned patients and 400,000 non- opioid patients to balance the dataset. Health Facts included 110,000 patients (10,000 positive, 100,000 negative), ensuring at least one hospital visit with documented clinical event records for improved analysis. Key predictive features included diagnosis codes, procedure codes, medications, clinical events, and demographic data. Diagnosis codes, essential for predict- ing future opioid overdoses, were filtered to remove direct opioid-related codes. Procedure codes were retrieved for both datasets, while medication data (NDC codes) were specific to Health Facts. Only the most relevant 10% of medication codes were retained, considering dosage and duration. Clinical events, available only in Health Facts, covered symptoms, pain levels, smoke history, and vitals. Since 79.21% hospitals included in the Health Facts recorded clinical events, they were valuable for prediction. Demographic information (age, gen- der, race) was included in both datasets, while payment sources (Medicare, Medicaid, etc.), relevant to socioeconomic status, were incorporated into SPARCS.This feature selection ensured a focused, efficient model for predicting opioid overdose risk.

D. Feature Selection and Normalization

To optimize prediction and training efficiency, we filtered features by selecting only the top 10 percent of diagnosis, procedure, and medication codes based on their frequency in opioid poisoning cases. This reduced the feature space while retaining relevant predictors. Health Facts includes a broader range of features compared to SPARCS, and the final selected features are summarized in Table V.

TABLE V: Overview of Selected Features for Prediction in Study Datasets

Datasets	**Category**	**Features**	**Description**
SPARCS	Diagnosis	2000	ICD-9, ICD-10 Codes
	Procedure	2000	ICD, CPT Codes
	Demographics	4	Race, gender, age, payment
Health Facts	Diagnosis	2000	ICD-9, ICD-10 Codes
	Procedure	1000	ICD, CPT Codes
	Demographics	3	Race, gender, age
	Clinical Events	900	500 events, 400 numeric values
	Medication	4500	1500 NDC Codes, dosage, duration

Ages were grouped into 0-5 years and then 10-year intervals to enhance training efficiency with minimal performance loss. One-Hot Encoding was applied to categorical features like

race and payment methods. Figure 1 illustrates this encod- ing for payment sources, where Patient 1 receives the code [1,1,0,0,0,0]. We represented diagnosis, procedure, medication codes, and clinical events using a binary format (1 if present, 0 if absent). Numeric features, such as medication dosage, duration, blood pressure, height, and pain score, were recorded in appropriate units (e.g., height in cm, blood pressure in mmHg). The feature preprocessing process is illustrated in Figure 2.

TABLE VI: One-Hot Encoding for Payments

Patient	**Self-Pay**	**Medicare**	**Medicaid**	**Insurance**
P1	1	0	0	0
P2	0	0	0	1
P3	1	0	1	0

This study seeks to predict opioid poisoning using EHR data through machine learning (ML) and deep learning (DL) models.

Traditional ML approaches, such as random forest, decision tree, and logistic regression, were used with opti- mized parameters. To address class imbalance, higher weights were assigned to positive cases. For DL, we developed and implemented a fully connected neural network with three layers, dropout regularization (20%), and ReLU activation for hidden layers. A sigmoid activation function was employed for binary classification. The model framework is illustrated in Figure 3. This study predicts opioid poisoning using EHR data with machine learning (ML) and deep learning (DL) approaches Traditional ML methods include random forest, decision tree, and logistic regression models with optimized parameters. To handle class imbalance, higher weights were assigned to positive cases. For DL, we used a fully connected neural network with three layers, dropout (20 percent), ReLU activation, and a sigmoid output for classification.

E. Prediction Methods

This study aims to predict opioid poisoning in patients using EHR data by applying both traditional machine learning and deep learning techniques. Traditional models such as decision trees and random forests, and logistic regression have proven effective for health data analytics. In our experiments, we prioritized recall over precision to ensure high-risk patients

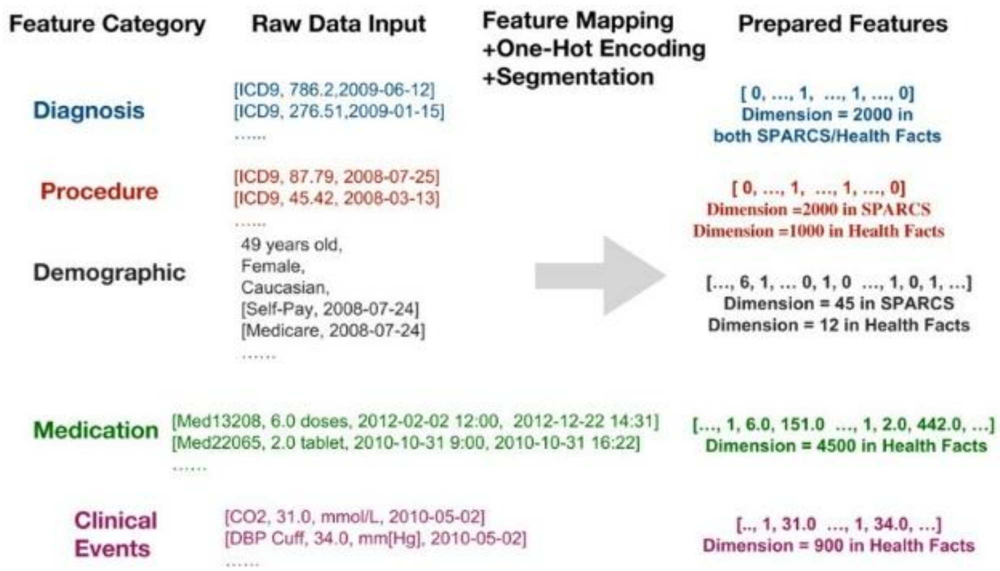

Fig. 1. Example of feature preprocessing for the prediction models.

were identified. Logistic regression applied L2 regularization, while decision trees and random forests used Gini impurity for data splitting. To address class imbalance, we assign a higher weight to positive cases, improving the F1 score. Figure 1 illustrates the sample preprocessing done for the prediction models.

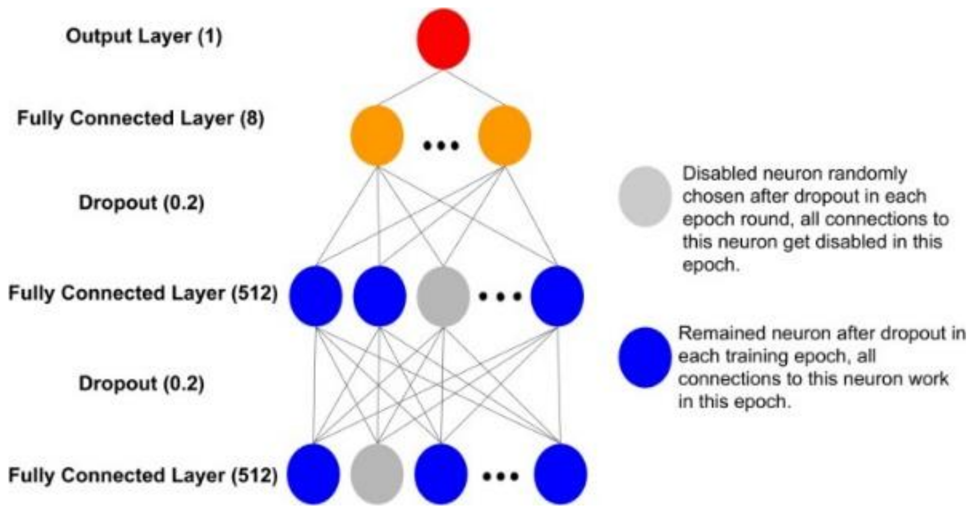

Fig. 2. Neural Network Framework for the Prediction Model

For deep learning, we implemented a fully connected neural network with three layers, two dropout layers, and a final output layer for binary classification. ReLU was used in hidden layers to prevent vanishing gradients, a sigmoid function was applied in the output layer, and the model optimized binary cross-entropy loss using the Adam algorithm for better effi- ciency. We conducted experiments using Python (TensorFlow, Keras, Scikit-learn) on an NVIDIA Tesla V100 GPU, ensuring robust and efficient implementation.

3. CONCLUSION

The opioid epidemic has emerged as a public health emer- gency in the United States. Identifying high-risk patients for opioid overdose can enable smarter and safer clinical decision- making, ultimately improving pain management practices. Our research on machine learning-based predictive models for opioid overdose detection has shown promising results using both claims data and comprehensive EHR data. These findings suggest that an AI-driven approach, when integrated into clinical settings, can enable automated and highly accurate predictions, offering valuable support to healthcare providers in addressing the opioid crisis.

REFERENCES

[1] Han B, Compton WM, Blanco C, Crane E, Lee J, Jones CM. Prescription opioid use, misuse, and use disorders in US adults: 2015 National Survey on Drug Use and Health. Annals of Internal Medicine. 2017 Sep 5;167(5):293–301.

[2] Rudd RA. Increases in drug and opioid-involved overdose deaths—United States, 2010–2015. MMWR Morbidity and Mortality Weekly Report. 2016:65.

[3] Scholl L, Seth P, Kariisa M, Wilson N, Baldwin G. Drug and opioid- involved overdose deaths—United States, 2013–2017. Morbidity and Mortality Weekly Report. 2019 Jan 4;67(5152):1419.

[4] Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. 2016;315(15):1624–45.

[5] Bruneau J, Ahamad K, Goyer ME` , Poulin G, Selby P, Fischer B, Wild TC, Wood E. Management of opioid use disorders: a national clinical practice guideline. Canadian Medical Association Journal. 2018 Mar 5;190(9):E247–57.

[6] Henry J, Pylypchuk Y, Searcy T, Patel V. Adoption of electronic health record systems among US non-federal acute care hospitals: 2008-2015. ONC Data Brief. 2016 May;35:1–9.

[7] Cheng Y, Wang F, Zhang P, Hu J, editors. Risk prediction with electronic health records: A deep learning approach. Proceedings of the 2016 SIAM International Conference on Data Mining; 2016: SIAM.

[8] Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities, and challenges. 2017;19(6):1236–46.

[9] Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. 2018;22(5):1589–604.

[10] Wang F, Casalino LP, Khullar D. Deep Learning in Medicine—Promise, Progress, and Challenges. 2018.

[11] Statewide Planning Research Cooperative System. [Internet]. Health.ny.gov. [cited 7 March 2019]. Available from: https://www.health.ny.gov/statistics/sparcs/

[12] Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. 2018;1(1):18.

[13] Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. BMC Medical Informatics and Decision Making. 2018 Dec;18(4):122.

[14] Lipton ZC, Kale DC, Elkan C, Wetzel R. Learning to diagnose with LSTM recurrent neural networks. 2015 Nov 11; arXiv preprint arXiv:1511.03677.

[15] Esteban C, Staeck O, Baier S, Yang Y, Tresp V, editors. Predicting clinical events by combining static and dynamic information using recurrent neural networks. 2016 IEEE International Conference on Healthcare Informatics (ICHI); 2016: IEEE.