Enhancing the National Early Warning Score 2 Using Observational Patient Data and Machine Learning.

Harry Abel
School of Medicine, Newcastle University, Newcastle upon Tyne, UK

Background

The National Early Warning System 2 (NEWS2) is a standardized clinical tool used across the NHS to monitor and detect early signs of patient deterioration based on vital signs such as heart rate, respiratory rate, and oxygen saturation. Early identification of deteriorating patients is critical to improving outcomes and reducing adverse events in hospital settings [1]. However, the predictive power of NEWS2 can be enhanced by leveraging machine learning to analyze complex patterns in patient data [2].

This work has the potential to support clinicians in making timely interventions, ultimately improving patient care and safety.

Aims

- This project aimed at using a machine learning model to accurately detect when a patient is deteriorating.
- To determine which additional variables are important in predicting a patient's next NEWS2 score.
- To test and validate the Model's performance using AUC, Accuracy and F1-score.

Methodology

This was a retrospective observation study using pseudonymised patient data collected from Newcastle upon Tyne Hospitals NHS foundation trust from Dec 2018 – June 2024 This resulted in 45,161 unique patients in this period [3].

NEWS2 scores 0-3 made up ~90% of the data indicating a big class imbalance. To overcome this, it was decided to categorize in 4 categories based on clinical need.

Methodology

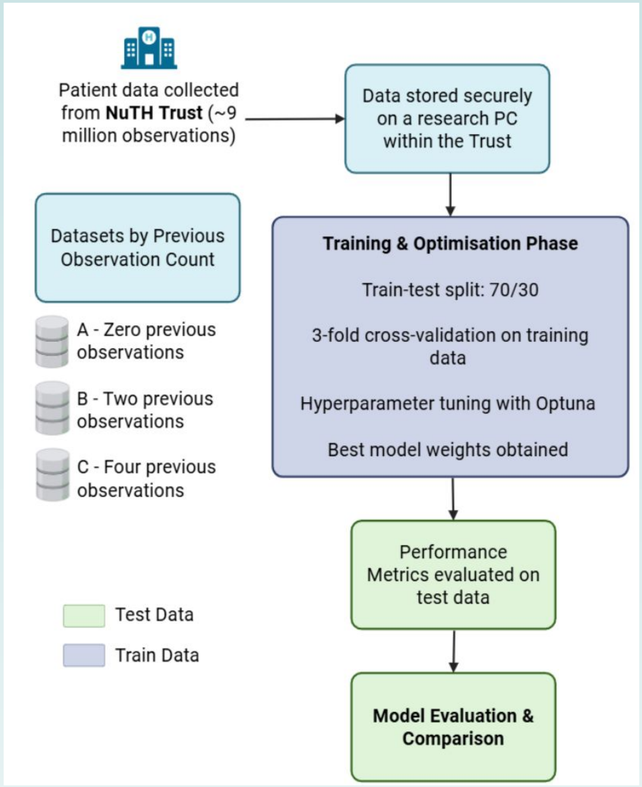

Figure 1. Schematic overview of the methodology, showing data collection from NuTH Trust, secure storage, dataset grouping, training and optimisation with cross-validation and hyperparameter tuning, and final model evaluation.

Methodology

An XGBoost model was trained within Python using the following packages: pandas, numpy, scikit-learn, XGBoost and Optuna.

Results

The most important variables are able to be determined from the model weights and the most important features for each model varied slightly but the **top features included**:

- ***NEWS2Score***
- ***AVPU***
- ***Respiratory Rate***
- ***Oxygen Device***
- ***Oxygen Saturation*.**

The other important features for the models that included previous observations were the previous NEWS2 Scores.

Also shown in the results is the accuracy of the respective models of 0.142, 0.454, and **0.646** with the highest accuracy being the model with data that included the most previous observations.

Results

	Model Type (Accuracy)		
Feature Ranking	0 Previous Observations **(0.142)**	2 Previous Observations **(0.454)**	4 Previous Observations **(0.646)**
1	NEWS2	NEWS2_-1	NEWS2_-1
2	MaskCode	MaskCode	AVPU
3	NuthRiskRating	NEWS2_-2	NEWS2_-2
4	OxyenTherapy	AVPU	MaskCode
5	OxgenSaturation	RespirationRate	RespiratoryRate

Table 1 - Showing the top 5 most important features extracted from each XGBoost model.

Discussion

The data shown in Table 1 shows that patient deterioration can be determined from previous observational data. Table 1 also shows that an increase in the amount of previous observation data can increase the predictive power of the model which is a similar trend seen by *Choi et al [2].*

The feature importance also indicates that the previous NEWS2 score, AVPU score, and Mask Code are important features in predicting the next NEWS2 score of a patient.

Physiological parameter	Score 3	2	1	0	1	2	3
Respiration rate (per minute)	≤8		9–11	12–20		21–24	≥25
SpO_2 Scale 1 (%)	≤91	92–93	94–95	≥96			
SpO_2 Scale 2 (%)	≤83	84–85	86–87	88–92 ≥93 on air	93–94 on oxygen	95–96 on oxygen	≥97 on oxygen
Air or oxygen?		Oxygen		Air			
Systolic blood pressure (mmHg)	≤90	91–100	101–110	111–219			≥220
Pulse (per minute)	≤40		41–50	51–90	91–110	111–130	≥131
Consciousness				Alert			CVPU
Temperature (°C)	≤35.0		35.1–36.0	36.1–38.0	38.1–39.0	≥39.1	

Figure 2. National Early Warning System 2 scoring systems and associated observational features. Image obtained from [4].

Conclusion

The performance of the model has shown that acute deterioration can be predicted using previous observational data. Additionally found what variables are important in predicting patient deterioration.

This work has the potential to support clinicians in making timely interventions, ultimately improving patient care and safety.

References

1 - Kemp K, Alakare J, Harjola V-P, et al. National Early Warning Score 2 (NEWS2) and 3-level triage scale as risk predictors in frail older adults in the emergency department. BMC Emerg Med. 2020;20:1–9. doi: 10.1186/s12873-020-00379-y

2 - Choi A, Choi SY, Chung K, et al. Development of a machine learning-based clinical decision support system to predict clinical deterioration in patients visiting the emergency department. Sci Rep. 2023;13:8561. doi: 10.1038/s41598-023-35617-3

3 - Cruz Rivera S, Liu X, Chan A-W, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nat Med. 2020;26:1351–63. doi: 10.1038/s41591-020-1037-7

4. Indd, N. (2017). National Early Warning Score (NEWS) 2 Standardising the assessment of acute-illness severity in the NHS. www.rcplondon.ac.uk

Acknowledgement

With thank to the knowledge and support of Professor Edward Meinert, Dr Cen Cong, Aditya Singh, Hanan Albalawi , Uthpala Molligoda, Peta Le Roux, Dr Lynsey Threlfall, and Dr Chris Plummer.

NIHR | Newcastle Biomedical Research Centre

Newcastle University