

# HAD7001-S25 Datathon 3 Submission by Team 2

Members: Cal Robinson, Daniel J Lee, Marya Zaidi

## Introduction

One-third of ICU-admitted adults die during hospitalization. Predicting ICU mortality risk is critical for early intervention and efficient resource allocation. Given the time-sensitive nature of ICU care, prediction models should use data from the first 24 hours of admission. This study develops a hospital mortality risk prediction model based on this early data. An ability to predict which patients admitted to the ICU are at increased risk of mortality is critical. Accurate predictions can facilitate early escalation strategies and interventions to reduce the risk of mortality among high-risk patients. It can also be used to allocate and manage critical care resources effectively at a hospital or health system level. Given the time-sensitive nature of ICU care, prediction models should leverage data available from within the first 24 hours of ICU admission. Therefore, we aimed to develop a hospital mortality risk prediction model using data from the first 24 hours of ICU admission among hospitalized adults.

## Methods

### Data engineering

We used data from adults ( $\geq 16$  years old) admitted to the ICU and enrolled in the MIT GOSSIS community initiative. We then used clinical expertise to determine clinical relevance of available features in the full dataset. After review, we included the following demographic factors (age, BMI, ethnicity, sex), hospitalization factors (ICU readmission, ICU type), clinical factors (GCS score, mechanical ventilation, day 1 minimum O2 saturation, day 1 maximum systolic blood pressure, day 1 minimum systolic blood pressure, day 1 maximum heart rate, day 1 maximum temperature, acute kidney injury, FiO2), laboratory measures (urea, creatinine, glucose, bicarbonate, INR, lactate), and patient comorbidities (immunosuppression, cancer, hepatic failure). All predictor features were collected from the first day of ICU admission.

We assessed missing data, outliers, and variable distributions using histograms and correlation matrices. The dataset was randomly split into training (70%) and testing (30%) sets. Missing data was imputed separately using multiple imputation (10 iterations), and predictor variables were standardized. We evaluated for missing data, outlier values, and used histograms and QQ plots to visually explore the distribution of the data. Correlation between features and outcomes of interest was estimated and visualized using a correlation matrix. The dataset was randomly split into training (70%) and testing (30%) sets. We checked for class imbalance across the full dataset and split training/test datasets. Missing data was imputed by multiple imputation (10 iterations) in the split training/testing datasets, separately. All predictor variables in the training and test datasets were rescaled and standardized to have a mean of 0 and variance of 1.

### Model Development and Evaluation

We trained a neural network model to predict in-hospital mortality using the first 24 hours of ICU admission data. Feature selection was performed, and model hyperparameters were optimized to improve performance. Multiple evaluation metrics, including accuracy, AUC-ROC, precision, recall, and f1-score, were used to assess model performance. The training process included 21 epochs before early stopping was triggered due to a lack of improvement in validation loss. The final model achieved a training accuracy of 0.921 and validation accuracy of 0.920. The model also demonstrated strong performance based on the AUC-ROC score of 0.85, indicating a good balance between sensitivity and specificity.

Precision-recall metrics and the confusion matrix revealed high precision (0.93) and recall (0.99) for survival (class 0), but lower recall (0.24) for mortality (class 1), indicating challenges in identifying high-risk patients. While the model had high precision (0.93) and recall (0.99) for patients who survived (class 0), the recall for patients who died (class 1) was much lower at 0.24. This suggests that the model struggled to correctly identify all high-risk patients. Class imbalance was evident in the dataset, with significantly fewer patients in the mortality group (8.6%) compared to the survival group (91.4%). This imbalance likely contributed to the reduced recall for class 1 predictions. Additional models, including XGBoost, were evaluated. The best-performing XGBoost model achieved an accuracy of 0.923, an AUC-ROC of 0.85, and a

classification report indicating improved precision for mortality predictions (0.70) but still limited recall (0.21). The confusion matrix for XGBoost also showed a high false negative rate, reinforcing the challenge of identifying at-risk patients.

**Outcome Definition:** Primary outcome of interest - In-hospital mortality

## Results

**Model Performance:** The neural network model showed robust overall classification ability with an accuracy of 0.92 and an AUC-ROC of 0.85. However, it demonstrated limited effectiveness in predicting mortality cases, as indicated by the low recall for class 1 (0.24).

**Class Imbalance Effect:** The dataset exhibited substantial class imbalance, with only 8.6% of patients classified as deceased. This led to a bias in model predictions, where survival cases were identified with high confidence, but mortality cases were frequently missed.

**Alternative Model Comparisons:** XGBoost provided a slightly better balance between precision and recall for mortality cases, improving precision (0.70) but still struggling with recall (0.21).

**Potential for Improvement:** To address performance disparities, future iterations should explore oversampling or undersampling techniques, cost-sensitive learning, or ensemble approaches to enhance mortality prediction.

**Clinical Implications:** The model effectively predicts survival but struggles with high-risk patient detection, necessitating further refinements before clinical deployment. Further refinements are required to enhance recall for mortality cases before considering integration into ICU decision-support systems.

## Discussion

Our neural network model shows robust performance with an accuracy of 0.92 and an AUC-ROC score of 0.85. This indicates strong classification performance of the model. However, the precision-recall analysis indicates a performance disparity between classes. Specifically, while group 0 (survival) achieved a high precision (0.93) and recall (0.99), group 1 (death) exhibited lower recall (0.24), suggesting a challenge in correctly identifying patients at risk of mortality. In other words, the model is highly confident at predicting survival, but at the cost of failing to flag critical cases which may lead to mortality. This could lead to missed opportunities for early medical intervention, potentially worsening patient outcomes. Therefore, our model is not yet ready to be deployed to aid in clinical decision-making.

The limitation of our model may stem from the class imbalance. This may be addressed in the future iterations through techniques like oversampling or undersampling. In addition, hyperparameter tuning for more complex models and cross validation could be explored to optimize results.

## Individual contributions

All team members separately performed data engineering and conducted the analysis. Results from each analysis were compared to ensure consistency. Cal was responsible for compiling code from each team member, conducting the final analysis. Cal was responsible for writing the introduction and data engineering process sections. Marya was responsible for writing the analysis and findings sections, hosting the analysis file on GitHub, and preparing the presentation slides. Daniel was responsible for writing the discussion section, reviewing the compiled code, and revising the final submission.

## Code and presentation

GitHub link: [https://github.com/Marya1ZZ/UofT\\_ML\\_Datathon3/tree/main](https://github.com/Marya1ZZ/UofT_ML_Datathon3/tree/main)

Presentation link: [https://docs.google.com/presentation/d/1iu3KR08GzqLlIWHtW-CqYPlaqC5IDSJCGxml8cJ\\_hLc/edit?usp=sharing](https://docs.google.com/presentation/d/1iu3KR08GzqLlIWHtW-CqYPlaqC5IDSJCGxml8cJ_hLc/edit?usp=sharing)