**Datathon #4:**

**Utilizing Neural Networks to Predict Early Mortality in Intensive Care Unit Patients**

**Hongyan Chen**

**Tamneet Tiwana**

**Omid Jafarinezhad (Team 15)**

**APPLIED MACHINE LEARNING FOR HEALTH DATA**

**November 14, 2023**

**Course Director**

**Zahra Shakeri**



**Introduction**

Mortality rates in the Intensive Care Units (ICUs) range from 9% to 61% worldwide, thus representing a critical challenge that negatively impacts families and burdens the healthcare system(1). The early prediction of mortality in the ICU helps with clinical decisions, optimizing resource allocation, and healthcare costs, and preparing families for potential negative events. Neural networks have been used in research to predict mortality. As such, Avanti et al. applied a deep neural network model that was able to identify patients at an elevated risk of mortality within the next three to twelve months(2).

Clinical factors such as oxygen saturation, acute respiratory failure, body temperature, BMI, and demographic factors including gender and age were strongly associated with poorer ICU survival(3-5). Similarly, health conditions such as intubation, mechanical ventilation, liver cirrhosis, hepatic failure, glycemic levels, diabetes mellitus (DM), immunosuppression, and diagnosis of leukemia, Acquired immunodeficiency syndrome (AIDS), or lymphoma increase the chances of mortality(4,6-8).

By leveraging a privacy-certified dataset from MIT’s Global Open Source Severity of Illness Score (GOSSIS) initiative, we aim to assess whether clinical aspects as mentioned can predict mortality in the first 24 hours following admission to ICU. To do so, we employed a Neural Network model to predict mortality in the first 24 hours following admission to ICU. Then, using the machine learning pipeline process, we will outline the data pre-processing and prediction techniques and evaluate the performance of the Neural Network model in predicting mortality in the ICU in the first 24 hours.

**Data-Engineering Process**

The dataset used for this analysis was provided in collaboration with MIT’s GOSSIS community initiative. From this dataset, we sought to explore the ability of a combination of the selected features of acute respiratory failure, age, gender, immunosuppression, glycemic levels, intubation, mechanical ventilation, hepatic failure, DM, cirrhosis of the liver, intubation, prevalence of AIDS, lymphoma, and leukemia, based on their clinical relevance, to identity the risk of mortality in the ICU in the first 24 hours. First, duplicate patient records were identified by their IDs, and earlier admission data were removed to ensure uniqueness. Missing values in critical variables were assessed by comparing descriptive statistics of records with and without missing data and checking for any impact of these missing values on the mortality outcome. The analysis indicated that the missing data were likely completely random, which justified using mean imputation for data replacement. Outliers of the predictor variables were examined by data points greater than three standard deviations. The distribution of variables was then explored through various graphical representations, including bar graphs, histograms, and pie charts to understand the feature-outcome relationship for mortality within 24 hours. There were 7915 individuals who passed away within 24 hours of admission into the ICU (8.6%), while 83,798 (91.4%) survived this period, indicating a significant outcome imbalance. Upsampling the minority class may be necessary in model-building to prevent biased results.

The data was randomly split into an 80-20% train-test section, with random oversampling applied to the training subset to generate a balanced distribution of the 24-hour mortality outcomes. The validation set was created by further splitting the training set with 80% of the data used for model training and 20% of the data used for validation, using the hold-out validation method.

**Data analysis**

Each model trained over 30 epochs with a batch size of 128 and a learning rate of 0.0005 using the Adam optimizer. This was employed due to its effectiveness in quickly adapting the learning rate and finding a minimum with minimal computational resources. Additionally, Binary Cross-Entropy with Logits Loss was used as the loss function, due to its effectiveness in binary classification.

The first neural network model was designed as a baseline model with a simple structure for predictions. Next, the second model incorporated dropout layers for regularization and introduced additional hidden layers to analyze the complex patterns and reduce overfitting. The third model employs a five-layer architecture with ReLU activations and dropout regularization. This model concatenated the outputs from previous layers to feed into subsequent ones, thereby enhancing its ability to extract and learn from a comprehensive set of data features. The model with the best performance (e.g., ROC-AUC and accuracy) was selected.

After training, the model's performance was evaluated on the validation set to assess its accuracy, precision, recall, F1 score, and ROC-AUC score. These metrics provided a comprehensive evaluation of the model's ability to generalize to unseen data and its potential utility in a clinical setting.

**Findings**

Within the dataset, there are 91713 observations of ICU visits at various hospitals, with 186 variables. In the dataset, missing values were most prevalent in the 'glucose\_apache' variable, with 11,036 missing entries, followed by 'BMI' and 'h1\_spo2\_min' at 3,429 and 4,185, respectively. Less frequent missing values in clinical conditions such as 'arf\_apache,' 'immunosuppression,' and 'ventilated\_apache,' each with 715 missing records, while 'age' data were absent in 4,228 cases. As such, mean imputation was conducted for data replacement.

Modelling neural networks showed some significant insights into mortality prediction within 24 hours of admission into the ICU. The initial model demonstrated an overall moderate performance with an accuracy of 0.71, precision of 0.72, recall of 0.69, and a ROC-AUC score of 0.71.

Similarly, the second model demonstrated moderate accuracy, indicating that the model is able to correctly predict mortality 71.70% of the time. The moderate recall (73.99%) suggests that the model can adequately predict mortality (moderate sensitivity) and the moderate precision (70.76%) indicates the model yields false positives (moderate specificity).

However, the final model demonstrated a high accuracy, indicating that the model is able to correctly predict mortality 80.33% of the time. The high recall (79.83%) suggests that the model can correctly predict mortality within 24 hours of ICU admission (high sensitivity) and the high precision (80.63%) indicates the model yields false positives (high specificity).

**Conclusion**

In summary, the use of a large and diverse dataset covering multiple hospitals worldwide strengthens the robustness and versatility of our final model to predict 24-hour mortality from ICU admission. However, due to the final model being intricate and requiring higher computational power, the implementation may pose a challenge in various healthcare settings. The consistent increase in all of the performance metrics, specifically the ROC-AUC scores across the models from basic to more advanced architectures highlights the value of utilizing deep learning techniques to capture complex clinical data relationships. This work allows healthcare professionals to identify and prioritize at-risk patients within the crucial first 24 hours of ICU admission based on pertinent demographic, physiological, and health factors.

**Team and Individual Contributions:** HC, TT, and OJ all contributed to the project design. All members contributed to coding, writing the report and presentation.

**Github:**[**hongyan627/Datathon4 (github.com)**](https://github.com/hongyan627/Datathon4)

**Presentation:**<https://docs.google.com/presentation/d/1DSDhgzs3SbXSqAkt8Ihk0CG0kOAMu_NjSHrCxBOI1Lo/edit?usp=sharing>

**References**

1. Demass TB, Guadie AG, Mengistu TB, Belay ZA, Melese AA, Berneh AA, Mihret LG, Wagaye FE, Bantie GM. The magnitude of mortality and its predictors among adult patients admitted to the Intensive care unit in Amhara Regional State, Northwest Ethiopia. Scientific Reports. 2023 Jul 25;13(1):12010.
2. [#] 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):55-64.
3. Barfod C, Lauritzen MM, Danker JK, Sölétormos G, Forberg JL, Berlac PA, Lippert F, Lundstrøm LH, Antonsen K, Lange KH. Abnormal vital signs are strong predictors for intensive care unit admission and in-hospital mortality in adults triaged in the emergency department-a prospective cohort study. Scandinavian journal of trauma, resuscitation and emergency medicine. 2012 Dec;20(1):1-0.
4. Goulenok C, Monchi M, Chiche JD, Mira JP, Dhainaut JF, Cariou A. Influence of overweight on ICU mortality: a prospective study. Chest. 2004 Apr 1;125(4):1441-5.
5. Soares Pinheiro FG, Santana Santos E, Barreto ID, Weiss C, Vaez AC, Oliveira JC, Melo MS, Silva FA. Mortality predictors and associated factors in patients in the intensive care unit: a cross-sectional study. Critical Care Research and Practice. 2020 Aug 1;2020.
6. Simpson A, Puxty K, McLoone P, Quasim T, Sloan B, Morrison DS. Comorbidity and survival after admission to the intensive care unit: A population-based study of 41,230 patients. Journal of the Intensive Care Society. 2021 May;22(2):143-51.
7. R. Brunner, G. Adelsmayr, H. Herkner, C. Madl, and U. Holzinger, “Glycemic variability and glucose complexity in critically ill patients: a retrospective analysis of continuous glucose monitoring data,” *Critical Care*, vol. 16, no. 5, p. R175, 2012.
8. Garnacho-Montero J, León-Moya C, Gutiérrez-Pizarraya A, Arenzana-Seisdedos A, Vidaur L, Guerrero JE, Gordón M, Martín-Loeches I, Rodriguez A, on Behalf GETGAG Study Group. Clinical characteristics, evolution, and treatment-related risk factors for mortality among immunosuppressed patients with influenza A (H1N1) virus admitted to the intensive care unit. Journal of critical care. 2018 Dec 1;48:172-7.