Mortality Prediction with Artificial Neural Network Method

[1]Dwight J.O Hutagalung , [2]Georgius Diko Parulian, [3]M Adli Rahmat Solihin, [4]M Alanna Daffa R.

[1234]Program Studi S1 Ilmu Aktuaria, Universitas Indonesia

[1234]Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indonesia

[1234]Jl. Prof. DR. Sudjono D. Pusponegoro, Universitas Indonesia, Pondok Cina, 16425

e-mail:

[1]*dwight.johutagalung@ui.ac.id*

[2]*georgius.diko@ui.ac.id*

[3]*muhammad.adli01@ui.ac.id*

[4]*muhammad.alanna@ui.ac.id*

Correspondence cell phone number: [1]081318809590 , [2]082111610296, [3]081385325445, [4]082169110155

ABSTRACT

The changing times and the complexity of factors influencing human mortality rates demand a new approach in mortality prediction. This article discusses the importance of using Artificial Neural Network (ANN) in mortality modeling as an effort to enhance prediction accuracy and health risk management. In the context of dynamic changes, the use of advanced technology such as ANN provides opportunities to integrate clinical and demographic variables, offering profound insights into factors affecting mortality. This research not only focuses on statistics but also delves into in-depth information to support more personalized and timely health intervention planning.

This article conducts a literature review related to mortality modeling, highlighting various machine learning methods such as Decision Tree, Gradient Boosting, and Random Forest applied in previous studies. Furthermore, the article discusses recent models like LightGBM, Regression Tree Boosting, and Deep Neural Network (DNN), showing promising results in predicting mortality. Through critical analysis of the literature, this article proves that the use of ANN, as a representation of Deep Learning models, significantly contributes to improving mortality prediction accuracy.

The research methodology is outlined by detailing the use of data from Kaggle, pre-processing data steps, and data division for training and testing ANN models. Model evaluation is performed through several metrics, including accuracy, precision, recall, F1-score, and AUC-ROC. The evaluation results provide an overview of the ANN model's performance, with an accuracy of 77%, precision of 84%, recall of 89%, F1-score of 86%, and AUC-ROC of 63%. This study makes a significant contribution to enhancing understanding of ANN implementation in mortality prediction, with significant implications for health risk management and medical decision-making.

Keywords: *Mortality; Artificial Neural Network (ANN); Prediction; Classification*

A. INTRODUCTION

The inevitable passage of time leaves humanity in a state of uncertainty. Evolving technologies and increasingly varied lifestyles perpetuate this uncertainty, impacting various aspects of human life, including mortality rates. Mortality rate serves as a metric to depict the number of individuals who perish within a population during a specific time period. Certainly, contemporary human mortality rates will fluctuate with changing lifestyles and environmental conditions. Predicting current mortality rates using existing methods proves challenging due to historical data becoming irrelevant to the present circumstances. However, mortality rates are crucial indicators in epidemiology, health statistics, demography, and social sciences.

In this era of dynamic change, mortality prediction utilizing Artificial Neural Network (ANN) has become a primary focus of research to enhance understanding and health risk management. Mortality, or the rate of death, not only serves as a population health indicator but also constitutes an integral element in quality medical decision-making. Within this framework, the use of advanced technology, such as ANN, presents new opportunities to improve prediction accuracy and, consequently, enhance the overall effectiveness of healthcare systems.

The significance of this research cannot be overlooked. By implementing Deep Learning models like ANN, we can effectively integrate various clinical and demographic variables, enabling us to gain profound insights into factors influencing mortality. This extends beyond mere statistics, encompassing a range of information that aids in planning more personalized and timely health interventions.

The outcomes of this research have significant implications in several domains. Firstly, improved prediction accuracy can assist healthcare service providers in planning resource allocation more efficiently. Understanding mortality risk at the individual level allows for better personalization of care, ensuring each patient receives attention tailored to their needs.

Moreover, the integration of ANN into clinical decision support systems opens the door to substantial improvements in medical decision-making. Physicians and healthcare professionals can access more accurate and relevant information, helping reduce diagnostic errors and enhance efficiency in providing care.

B. LITERATURE REVIEW

Mortality Modeling

Mortality modeling is a statistical or mathematical analytical tool employed to depict the relationships between factors contributing to the death rate within a population. This model can take the form of mathematical equations or statistical formulations linking variables such as age, gender, health factors, and other risk factors to the death rate. The objective is to identify mortality patterns, comprehend trends, and predict how death rates will evolve in the future. Mortality models find applications in various fields, including insurance, demography, epidemiology, and health research, aiding in planning, decision-making, and risk analysis. Mortality models can range from simple ones, such as linear regression models linking age to death rates, to more complex ones like the Lee-Carter or Cairns-Blake-Dowd models employing sophisticated statistical techniques to depict mortality patterns. The choice of the model depends on the available data and research objectives.

Current Developments

Levantesi and Pizzorusso (2019) applied machine learning techniques to assist in mortality modeling using Decision Tree, Gradient Boosting, and Random Forest. Random Forest emerged as the most effective in mortality modeling, enhancing the fitting quality of Lee-Carter, Renshaw-Haberman, and Plat models in the Italian population. The Random Forest method significantly reduced the highest MAPE achieved by the Plat model, decreasing from 25.81% to 4.79% after applying the Random Forest algorithm (from 22.34% to 4.49% for the female population).

Subsequently, LightGBM excelled in predicting sepsis patient mortality (Bao, C., et al., 2023). LightGBM achieved an average f1-score of 0.909483 during model training and 0.905742 during testing, outperforming other methods such as XGBoost, Multiple Layer Perception (MLP), Gradient Boosting Machine (GBM), Random Forest, Decision Tree, and Support Vector Machine (SVM). The study also identified SVM as the model with the lowest f1-score.

Regression Tree Boosting is another machine learning method applicable to mortality modeling. Deprez et al. (2017) utilized Regression Tree Boosting to analyze deficiencies in the Lee-Carter and Renshaw-Haberman models (a Lee-Carter model development). Besides detecting model shortcomings, Regression Tree Boosting can enhance the fitness of a model, providing a simple way to identify patterns in probability over time.

Furthermore, one of the Deep Learning models, DNN, was employed to predict mortality during a 1-year clinical follow-up after hospital discharge for ACS patients in Korea by Sherazi et al. (2020). The Precision performance of DNN surpassed machine learning models like GBM, GLM, Random Forest, as well as regression models like GRACE. However, GBM outperformed in AUC, Recall, Accuracy, and F-score compared to other methods. Despite equal AUC performance between GBM and DNN, GBM generally excelled in mortality prediction.

Artificial Neural Network Model

The Artificial Neural Network (ANN), commonly referred to as Neural Network (NN), is an algorithm motivated to construct a machine that mimics the brain's network system. Neural networks consist of a group of interconnected artificial cells. They are physical cellular systems capable of acquiring, storing information, and utilizing empirical knowledge. Similar to the human brain, ANN's knowledge is derived from the examples it encounters. In the human neural system, the learning process involves modifications to the synaptic connections between neurons. Similarly, ANN adjusts its structure based on incoming and outgoing information flowing through the network during the learning phase.

The data processing procedure in neural networks generally involves two main steps: learning and application. In the first step, a training database or historical price data is required to train the network. This dataset includes input vectors and known output vectors, with each input and output representing a node or neuron. Additionally, there are one or more hidden layers. The goal of the learning phase is to adjust the weight of connections between layers or nodes. After preparing the training samples, in an iterative approach, the samples are fed into the network, and its output is compared with the known output. If the result and the unknown output differ, changes to the weight of connections will continue until the difference is minimized. Upon achieving the desired convergence for the network in the learning process, a validation dataset is applied to the network for the validation step (Shahkarami A. et al., 2014).

C. RESEARCH METHODS

Data

The data utilized in this research was curated from the website Kaggle that is uploaded by the Laurence Gesman account on this link https://www.kaggle.com/code/laurencegesman/hospital-mortality-ml-data-analysis/input. The data employed is in the form of a .csv file, comprising a total of 60,087 data entries, detailed as 1,178 rows and 51 columns. This dataset contains information related to the medical records of a cohort of patients observed in a hospital, comprehensively monitoring their health conditions, and noting whether the patients survived or not during their hospitalization. The utilized data consists of 39 numerical variables and 12 categorical variables, with 50 variables serving as predictor variables and 1 categorical variable used as the target variable, namely the outcome variable describing the patient's life status, indicating whether the patient is still alive or has deceased.

Data Preprocessing

Data preprocessing involves a series of steps or techniques conducted on data before it is used as input for models or further analysis. The objective of data pre-processing is to prepare and cleanse the data so that it can be properly processed by Artificial Neural Network algorithms or models. Several data pre-processing steps undertaken in this study include data imputing, data splitting, and data standardization.

1. **Impute Data**

   Imputation stands out as the most judicious choice for handling missing data compared to discarding observations or variables containing missing values, considering the high cost and value of data. In this preprocessing phase, the researcher leverages sklearn.impute to fill variables with missing values, enabling their utilization in model development.

2. **Data Standardization**

   The primary objective of standardization is to bring all features to the same scale without distorting differences in value ranges. Given the diverse numeric scales across numerous variables, data standardization is imperative to support the model development process. In its execution, the researcher employs sklearn.preprocessing, importing MinMaxScaler and StandardScaler features.

3. **Data Splitting**

   Data splitting is a method of dividing data into two or more parts, forming subsets of data. Typically, data splitting separates two parts, with the first part used for evaluation or testing, and the other part used for model training. In the forthcoming deep learning model development, the available data is divided into training and testing sets, with an 80% composition for training data, 10% for validation data, and the remaining 10% for testing data.

Model Evaluation

The evaluation of the model will be compared based on three indicators: accuracy derived from the loss function, evaluation metrics, and the ROC-AUC method.

1. Loss Function

The loss function is a metric employed in machine learning and deep learning to gauge how well a model predicts target values. It measures the extent to which the model's prediction deviates from the desired actual value. In the context of deep learning, specifically in binary classification tasks, Binary Crossentropy is a commonly used loss function.

Binary Crossentropy, often referred to as Logarithmic Loss, is utilized when the classification task is binary, meaning there are two possible classes for each sample (typically positive and negative classes). This function compares the probability distribution generated by the model with the actual probability distribution of the target class.

2. Evaluation Metrics

Evaluation metrics are employed to assess the model after the data training stage. These metrics measure how well the model can generalize to new data and make accurate predictions. Evaluation metrics also compare various models or configurations to determine which performs best.

The properties or characteristics of evaluation metrics used to assess model performance include:

1. Confusion Matrix: A table summarizing the model's performance by breaking down predictions into true positive, true negative, false positive, and false negative instances.
2. Accuracy: Measures the overall correctness of the model's predictions, calculated as the ratio of correct predictions to total predictions.
3. Precision: Indicates the accuracy of positive predictions, calculated as the ratio of true positive predictions to the sum of true positive and false positive predictions.
4. Recall: Measures the model's ability to capture all positive instances, calculated as the ratio of true positive predictions to the sum of true positive and false negative predictions.
5. F1 Score: The harmonic mean of precision and recall, providing a balanced measure considering both false positive and false negative predictions.
6. Specificity: Measures the model's ability to correctly identify negative instances, calculated as the ratio of true negative predictions to the sum of true negative and false positive predictions.

3. ROC-AUC

The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) is a commonly used evaluation metric in the context of binary classification to measure the extent to which a model can distinguish between two classes. Let's delve into its explanation:

1. Receiver Operating Characteristic (ROC) Curve

   The ROC Curve is a graph that illustrates the performance of a classification model at various threshold levels to distinguish between positive and negative classes. On the X-axis, the ROC Curve shows the False Positive Rate, while on the Y-axis, it depicts the True Positive Rate or Sensitivity. Each point on the curve represents a specific threshold.

2. True Positive Rate (Sensitivity)

   Sensitivity represents the proportion of true positives identified by the model from the total number of actual positive cases. Measured as the ratio of True Positive (TP) divided by the total actual positives.

$$Sensitivity = \frac{TP}{TP+FN}$$

3. False Positive Rate (1-Specificity)

   It represents the proportion of actual negatives wrongly identified as positives by the model from the total number of actual negative cases. Measured as the ratio of False Positive (FP) divided by the total actual negatives.

$$False\ Positive\ Rate = \frac{FP}{TN+FP}$$

4. Area Under the Curve (AUC)

   AUC is the area under the ROC curve. It provides a concise measure of how well the model can distinguish between positive and negative classes. AUC values range between 0 and 1, where 1 indicates a perfect model (with 100% sensitivity and specificity), while 0.5 indicates model performance equivalent to random chance.

   The higher the AUC value, the better the model is at distinguishing between the two classes. If the AUC is around 0.5, it suggests that the model is no better than randomly predicting.

The higher the AUC value, the better the model is at distinguishing between the two classes. If the AUC is around 0.5, it suggests that the model is no better than randomly predicting.

AUC-ROC offers a comprehensive view of the classification model's performance at various thresholds and is especially useful when balancing sensitivity and specificity needs consideration. The closer the AUC-ROC value is to 1, the better the model's ability to differentiate between positive and negative classes.

D. RESULTS AND DISCUSSION

The performance of the constructed model is assessed using metrics such as accuracy to gauge the overall performance in classifying each class. However, a more in-depth examination of the model's performance is essential to evaluate how well the classification model operates according to the objectives of this research. The primary goal of model development is to predict the positive class or the minority value in the data, namely the target variable 'outcome' with a value of 1, indicating patient mortality.

Analyzing the model's performance in classifying specifically can be achieved by utilizing a confusion matrix, providing values to examine metrics such as accuracy, precision, recall, f1-score, and AUC-ROC.

The following is the confusion matrix obtained from the results of the testing phase of the built ANN model.

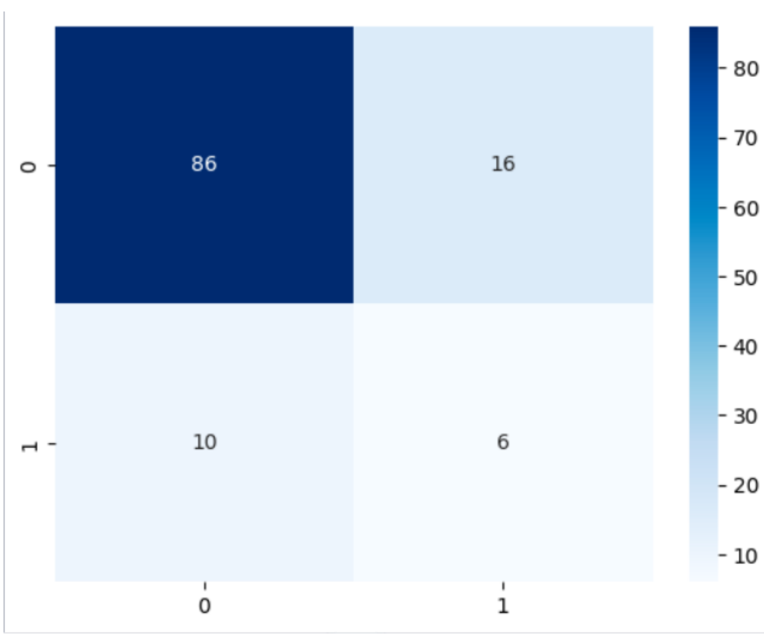

Figure 1. Confusion Matrix of ANN Model

From the confusion matrix, can be interpreted by the values of each entry are observed as follows:

- True Negatives (TN): 86 samples of the negative class were correctly predicted as the negative class.
- False Negatives (FN): 16 samples of the negative class were erroneously predicted as the positive class.
- False Positives (FP): 10 samples of the positive class were incorrectly predicted as the negative class.
- True Positives (TP): 6 samples of the positive class were accurately predicted as the positive class.

These values provide insights into the model's performance in correctly identifying instances of the positive and negative classes. The confusion matrix serves as a valuable tool to further analyze and refine the classification model's effectiveness.

The values obtained from the confusion matrix are utilized to compute metrics that can be employed to assess the performance of the constructed ANN model in classification. The results of these metric calculations are presented in the following table:

Tabel 1. Results of ANN Model Evaluation

Evaluation Metrics	***Value percentage(%)***
Accuracy	77%
Precision	84%
Recall	89%
F1-score	86%
AUC-ROC	63%

E. CONCLUSION

In the midst of the uncertainties posed by the changing times, this research underscores the significance of predicting human mortality rates through the utilization of Artificial Neural Network (ANN) technology. Lifestyle changes and environmental shifts demand an innovative approach to understanding and managing health risks, and ANN provides a promising solution. Through a literature review, this article explores various mortality models employed in diverse contexts, ranging from insurance to epidemiology. Current advancements, such as the application of machine learning and deep learning, have proven effective in enhancing mortality prediction accuracy.

The study notes that employing ANN models in mortality modeling can have a positive impact on several fronts. Improved prediction accuracy opens opportunities for healthcare providers to allocate resources efficiently and personalize care. The use of ANN in clinical decision support systems also paves the way for significant improvements in medical decision-making, helping reduce diagnostic errors and enhance treatment efficiency.

Through a detailed methodology, this article utilizes data from reputable sources and performs data pre-processing to ensure the reliability of the constructed ANN model. Model evaluation considers various metrics, including accuracy, precision, recall, f1-score, and AUC-ROC. The evaluation results indicate that the ANN model demonstrates good performance in classifying the positive class, which, in this context, represents patients with a mortality status. With an accuracy of 77%, precision of 84%, recall of 89%, f1-score of 86%, and AUC-ROC of 63%, the model provides valuable insights for planning more personalized and timely health interventions.

Overall, this research emphasizes that the use of advanced technologies such as ANN in mortality modeling is a relevant and beneficial step. With a deeper understanding of factors influencing mortality, we can enhance health risk management and deliver more effective care. Thus, this article makes a significant contribution to addressing health challenges in this ever-evolving era.

REFERENCES

Bao, C., et al. “Machine-Learning Models for Prediction of Sepsis Patients Mortality.” Medicina Intensiva (English Edition), vol. 47, no. 6, Nov. 2022, https://doi.org/10.1016/j.medine.2022.06.024. Accessed 5 Dec. 2022.

Deprez, Philippe, et al. “Machine Learning Techniques for Mortality Modeling.” European Actuarial Journal, vol. 7, no. 2, 6 June 2017, pp. 337–352, https://doi.org/10.1007/s13385-017-0152-4. Accessed 29 Sept. 2022.

Levantesi, Susanna, and Virginia Pizzorusso. “Application of Machine Learning to Mortality Modeling and Forecasting.” Risks, vol. 7, no. 1, 26 Feb. 2019, p. 26, https://doi.org/10.3390/risks7010026.

Sherazi, Syed Waseem Abbas, et al. “A Machine Learning–Based 1-Year Mortality Prediction Model after Hospital Discharge for Clinical Patients with Acute Coronary Syndrome.” Health Informatics Journal, vol. 26, no. 2, 30 Sept. 2019, pp. 1289–1304, https://doi.org/10.1177/1460458219871780. Accessed 24 Dec. 2020.

Ke, Guolin, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. 2017.