

The Area Under the Receiver Operating Characteristic Curve, often abbreviated as **AUC-ROC** or simply **AUC**, is a commonly used metric to evaluate the performance of **binary classification models**, particularly in **machine learning** and **statistics**. The **ROC** curve itself is a graphical representation of a model's ability to distinguish between two classes (usually a positive class and a negative class) across different threshold values. Here's an explanation of the AUC-ROC:

Receiver Operating Characteristic (ROC) Curve: The ROC curve is created by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold values for a binary classification model. TPR is also called Sensitivity or Recall, and it represents the proportion of actual positive cases correctly predicted as positive. FPR is the proportion of actual negative cases incorrectly predicted as positive. The ROC curve helps visualize a model's trade-off between sensitivity and specificity.

AUC-ROC: The AUC-ROC is a single scalar value that quantifies the overall performance of a binary classification model. It represents the area under the ROC curve. The AUC value ranges from 0 to 1, with higher values indicating better model performance. An AUC of 0.5 corresponds to a model that performs no better than random chance, while an AUC of 1 represents a perfect model.

An AUC of 0.5: The model's predictions are no better than random guessing. **An AUC between 0.5 and 1:** The model has some discriminatory power, with higher values indicating better performance. **An AUC of 1:** The model perfectly distinguishes between positive and negative cases. **Interpretation:** A higher AUC-ROC value indicates that the model is better at distinguishing between positive and negative cases. In practical terms, it suggests that as you vary the classification threshold, the model is consistently making better choices in terms of true positives and false positives. Therefore, a model with a higher AUC is generally considered more effective for binary classification tasks.

Use Cases: AUC-ROC is widely used in various domains, including healthcare (e.g., medical diagnosis), finance (e.g., credit scoring), and marketing (e.g., customer churn prediction). It is especially useful when evaluating models for imbalanced datasets, where one class significantly outnumbers the other.

Comparing Models: AUC-ROC provides a standardized way to compare different models for the same binary classification task. Models with higher AUC values are preferred because they exhibit better overall performance.

Limitations: While AUC-ROC is a valuable metric, it does not provide insight into other aspects of model performance, such as the specific threshold to use in a real-world application or the potential impact of false positives and false negatives in a particular context. For those purposes, other metrics like precision, recall, F1-score, and the confusion matrix may be more informative.

In summary, AUC-ROC is a widely used metric for evaluating the performance of binary classification models. It provides a concise summary of a model's ability to discriminate between positive and negative cases across different threshold values and is particularly useful for comparing and selecting models.

A receiver operating characteristic curve, i.e., ROC curve, is a graphical **plot** that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.

In a **ROC curve** the **true positive** rate (**Sensitivity**) is plotted in function of the false positive rate (100-Specificity) for different cut-off points of a parameter. ... The area under the ROC curve (**AUC**) is a measure of how well a parameter can distinguish between two diagnostic groups (diseased/normal).

In a [ROC curve](#) the [true positive](#) rate ([Sensitivity](#)) is plotted in function of the [false positive](#) rate (100-Specificity) for different cut-off points of a parameter. ... The [area under the ROC curve](#) ([AUC](#)) is a measure of how well a parameter can distinguish between two diagnostic groups (diseased/normal).

[Intraoperative neurophysiological monitoring](#) using [somatosensory evoked potentials](#) has been linked to a reduction in the [incidence](#) of [neurological deficits](#) during corrective surgery. Nonetheless, quantitative assessments of somatosensory evoked potential waveforms are often difficult to evaluate, because they are affected by anesthesia, injury, and noise.

Jorge et al., discuss a novel method that integrates somatosensory evoked potential signals into a single metric by calculating the area under the curve (AUC).

Thirty-two [Sprague Dawley rats](#) underwent a [laminectomy procedure](#) and were then randomly assigned to a control group or to receive a contusive [spinal cord injury](#) ranging from 100 to 200 kilodynes. Neurophysiological testing was completed at various points perioperatively and postoperatively. Somatosensory evoked potential traces obtained were processed and the AUC metric was calculated.

The AUC significantly decreased to 11% of its baseline value after impact and remained at 25% baseline after 1 hour for the 200-kdyn cohort. Postimpact, AUC for the control versus the 150-kdyn and 200-kdyn groups, and the 150-kdyn versus 200-kdyn groups were significantly higher ($P < 0.01$, $P < 0.001$, and $P < 0.05$, respectively). Across days, the only significant parameter accounting for AUC variability was impact force, $P < 0.0001$ (subject parameters and number of days were not significant).

The AUC metric can detect an iatrogenic contusive spinal cord injury immediately after its occurrence. Moreover, this metric can detect different [iatrogenic](#) injury impact force levels and identify injury in the postoperative period. The AUC integrates multiple Intraoperative neurophysiological monitoring measures into a single metric and thus has the potential to help clinicians and investigators evaluate spinal cord impact injury status ¹⁾.

The purpose of this study was to build a model of [machine learning](#) (ML) for the [prediction](#) of [mortality](#) in patients with isolated moderate and [severe traumatic brain injury](#) (TBI).

Hospitalized adult patients registered in the Trauma Registry System between January 2009 and December 2015 were enrolled in this study. Only patients with an [Abbreviated Injury Scale](#) (AIS) score ≥ 3 points related to head injuries were included in this study. A total of 1734 (1564 survival and 170 non-survival) and 325 (293 survival and 32 non-survival) patients were included in the training and test sets, respectively.

Using demographics and injury characteristics, as well as patient [laboratory](#) data, predictive tools (e.g., [logistic regression](#) [LR], [support vector machine](#) [SVM], [decision tree](#) [DT], [naive Bayes](#) [NB], and [artificial neural networks](#) [ANN]) were used to determine the [mortality](#) of individual patients. The predictive performance was evaluated by accuracy, [sensitivity](#), and [specificity](#), as well as by [area](#)

[under the curve](#) (AUC) measures of receiver operator characteristic curves. In the training set, all five ML models had a specificity of more than 90% and all ML models (except the NB) achieved an accuracy of more than 90%. Among them, the ANN had the highest sensitivity (80.59%) in mortality prediction. Regarding performance, the ANN had the highest AUC (0.968), followed by the LR (0.942), SVM (0.935), NB (0.908), and DT (0.872). In the test set, the ANN had the highest sensitivity (84.38%) in mortality prediction, followed by the SVM (65.63%), LR (59.38%), NB (59.38%), and DT (43.75%).

The ANN model provided the best prediction of mortality for patients with isolated moderate and severe TBI ²⁾.

¹⁾

Jorge A, Zhou J, Dixon EC, Hamilton KD, Balzer J, Thirumala P. Area Under the Curve of Somatosensory Evoked Potentials Detects Spinal Cord Injury. J Clin Neurophysiol. 2019 Jan 28. doi: 10.1097/WNP.0000000000000563. [Epub ahead of print] PubMed PMID: 30694945.

²⁾

Rau CS, Kuo PJ, Chien PC, Huang CY, Hsieh HY, Hsieh CH. Mortality prediction in patients with isolated moderate and severe traumatic brain injury using machine learning models. PLoS One. 2018 Nov 9;13(11):e0207192. doi: 10.1371/journal.pone.0207192. eCollection 2018. PubMed PMID: 30412613.

From:

<https://neurosurgerywiki.com/wiki/> - **Neurosurgery Wiki**

Permanent link:

https://neurosurgerywiki.com/wiki/doku.php?id=area_under_the_receiver_operating_characteristic_curve

Last update: **2025/05/13 02:10**

