

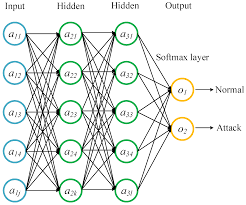
Confusion matrix in evaluation of intrusion detection systems

Cybercrimes are these crimes that are committed using computers and include mainly online financial fraud crimes, cyberterrorism, computer as a tool, computer as a target types of crimes that consist of phishing scams, identity theft, etc... These crimes may harm people’s security and even their financial health. Through the last decade, the number of this type of crime increased rapidly that conversely has led to development of some mitigation tools such as intrusion detection systems used to analyze network traffic in real time in order to detect malicious intention and take some actions. This article is an overview of confusion matrix in intrusion detection systems.

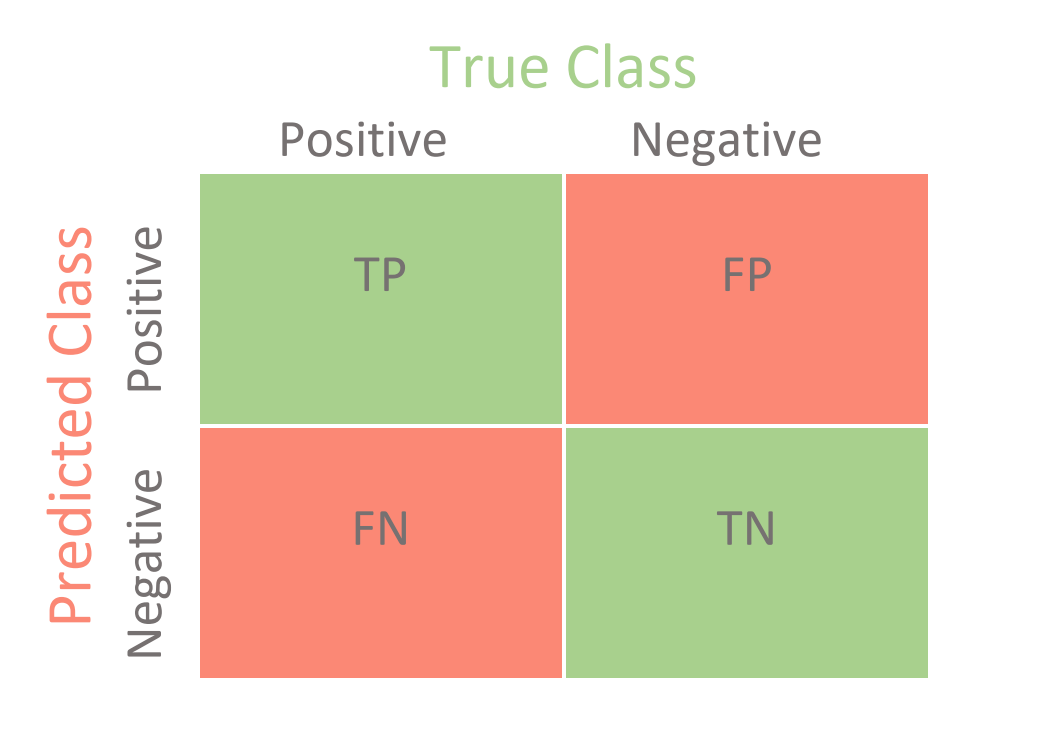
**What is confusion matrix?**

A confusion matrix is a type of table construct that plays a specific role in machine learning and related engineering. It helps to show the prediction and recall in a system where the values of the test data are known. The confusion matrix usually treats binary classification type of problems where the number of possible values of the results is two.

In binary classification type of problems such intrusion detection, the confusion matrix is a technique that helps to summarize the performance of classification of the model of the machine learning algorithm, classification which is based on a set of data where the final outcome is already known from a historical record.

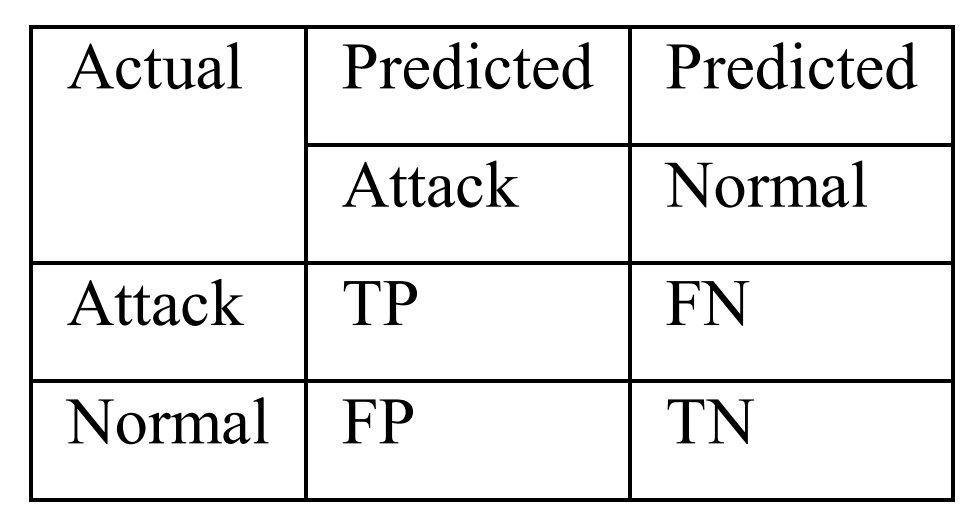


The confusion matrix consists a matrix of dimension equal to two where the columns are the outcome predicted values and the rows the actual values which are already known.



Meaning of confusion matrix table in intrusion detection

The confusion matrix of an intrusion detection system can be presented as follow:



It is to be recalled that the purpose of an IDS is to prevent against attack by detecting these attack that might be going in real time are labeled as positive values (label “Attack”), so, conversely the Normal values here means that there is no active attack going on, so they are negative (labeled “Normal”).

The followings are the possibilities to classify events and depicted in the above table:

* True positive (TP): Intrusions that are successfully detected by the IDS. So, the network is undergoing attacks and some decision should be taken to mitigate them.
* False positive (FP): Normal/non-intrusive behavior that is wrongly classified as intrusive by the IDS. The traffic is under normal situation
* True Negative (TN): Normal/non-intrusive behavior that is successfully labeled as normal/non-intrusive by the IDS. The model successfully classified attacks as attack
* False Negative (FN): Intrusions that are missed by the IDS, and classified as normal/non-intrusive. Here the model classified the attacks’ traffic as normal.

Let’s take a close look at some values of this matrix:

The confusion matrix gives an overview of the performance of the model at glance. Some important things to notice here are the FALSE cases:

The false positive: in real life and in classical model, this is usually also known as Type 1 Error and is known to be the most dangerous type of error in the machine learning because this case is classified not dangerous where the real danger is really happening. Here, according to our target (detect attack), this case instead falls under the type 2. This means that here it is false alarm, and there is no attack happening.

The false Negative: in classical type of problems, this case is usually known to be the less dangerous or the type 2 error. But here, it is to be recalled that the target of the model in the intrusion detection is to find intrusion as its’ name suggests, and from the above table, we can clearly see that this present case falls under the type 1 error, as the systems fails to detect some intrusion, which can be critical to a corporation, considering as example, an attack which aims to steal very sensitive informations as in case of a Bank.

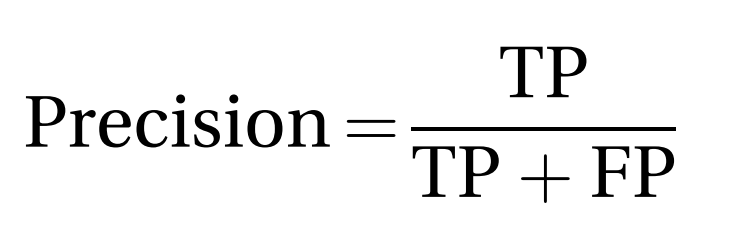
In real enterprises, the security operation team in companies are really preoccupied by these two errors and mostly by the type 1 error as at the end, team members can not fully rely on their infrastructure. This can be explained by the fact that no machine learning model can give absolutely right answer (zero percent of error). To mitigate these situations, they can perform proactivation (this is out of the scope this article).

The confusion matrix of IDS has led to some metrics:

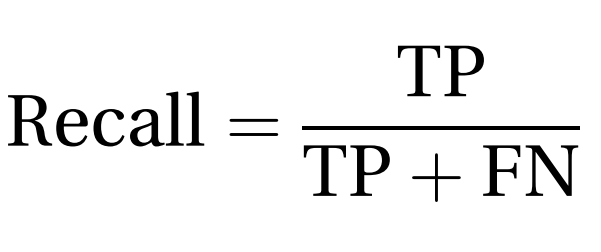
One of the most commonly used metrics while performing classification is accuracy. The accuracy of a model (through a confusion matrix) is calculated using the given formula below.

Precision and recall are widely used and popular metrics for classification.

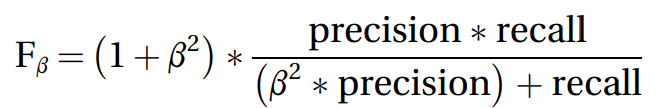
Precision shows how accurate the model is for predicting positive values. Thus, it measures the accuracy of a predicted positive outcome. It is also known as the positive predictive value. The formula to calcite the precision is :



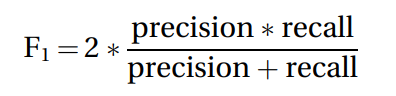
Recall is useful to measure the strength of a model to predict positive outcomes, and it is also known as the sensitivity of a model. The formula to calculate the recall is as follow:

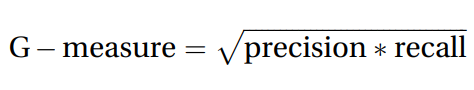


F-measure is also known as F-value which uses both precision score and recall score of a classifier. F-measure is another commonly used metric in classification settings. F-measure is calculated using a weighted harmonic mean between precision and recall. For the classification of positive instances, it helps to understand the tradeoff between correctness and coverage. The general formula for calculating F-measure is given below:



In the above formulation, the importance of each term can be provided using different values for b. Most commonly used value for b is 1, which is known as F-1 measure. G-measure is similar to that of F-measure, but it uses geometric mean instead of harmonic mean.





**Conclusion**

In short, the first thing not to forget is that no model of machine learning is perfectly accurate and we have to keep an eye on the faults of our model. To evaluate model performance many techniques have been developed, some of them based on the confusion matrix.

#worldrecordholder #training #internship #makingindiafutureready #summer #summertraining #python #machinelearning #docker #rightmentor #deepknowledge #linuxworld #vimaldaga #righteducation