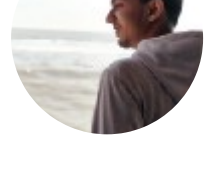


Understanding Confusion Matrix

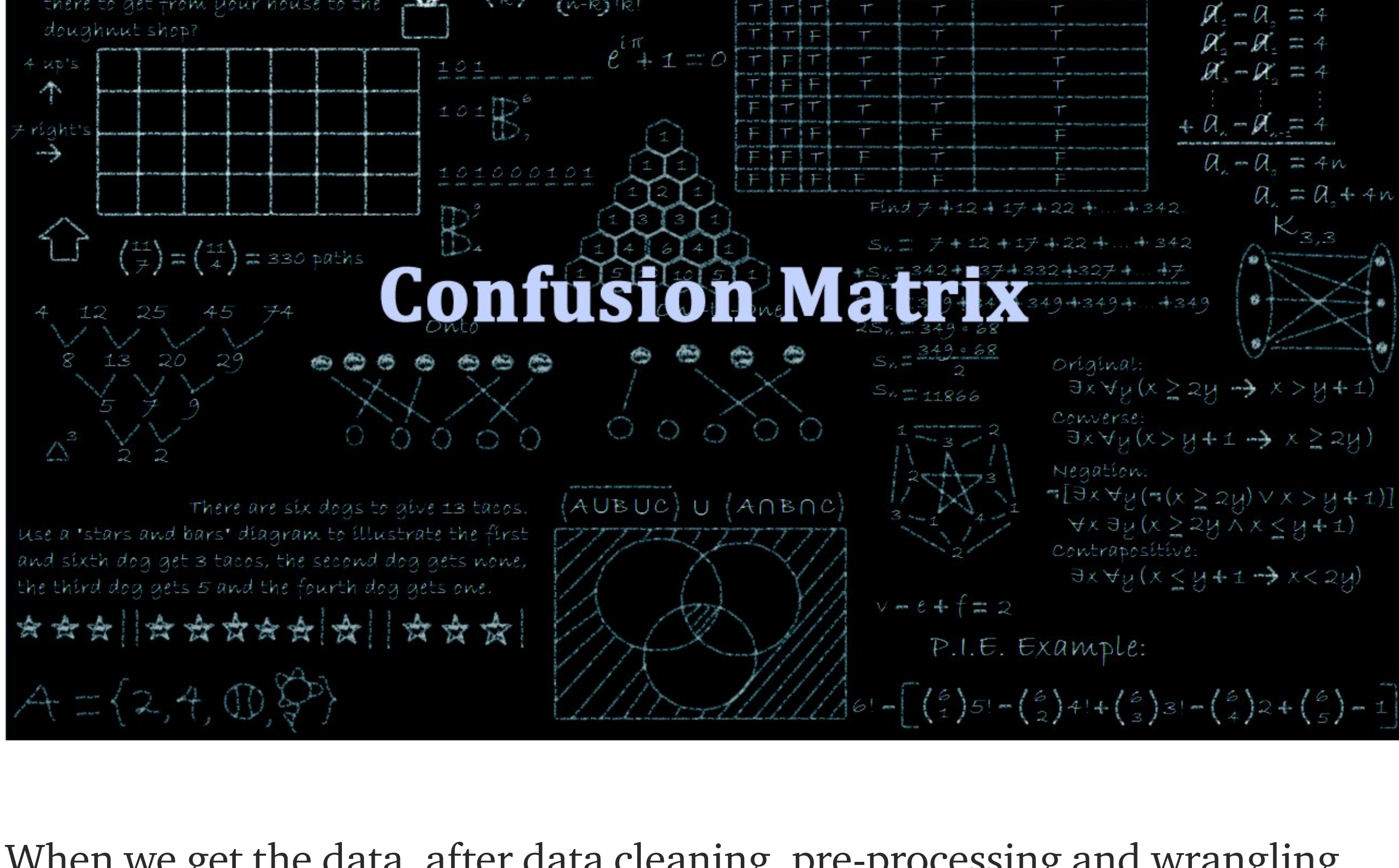


Sarang Narkhede

Follow

May 9, 2018 · 4 min read

🐦🌐📘📌🔗



When we get the data, after data cleaning, pre-processing and wrangling, the first step we do is to feed it to an outstanding model and of course, get output in probabilities. But hold on! How in the hell can we measure the effectiveness of our model. Better the effectiveness, better the performance and that’s exactly what we want. And it is where the Confusion matrix comes into the limelight. Confusion Matrix is a performance measurement for machine learning classification.

This blog aims to answer following questions:

1. What the confusion matrix is and why you need it?
2. How to calculate Confusion Matrix for a 2-class classification problem?

Today, let’s understand the confusion matrix once and for all.

What is Confusion Matrix and why you need it?

Well, it is a performance measurement for machine learning classification problem where output can be two or more classes. It is a table with 4 different combinations of predicted and actual values.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

It is extremely useful for measuring Recall, Precision, Specificity, Accuracy and most importantly AUC-ROC Curve.

Let’s understand TP, FP, FN, TN in terms of pregnancy analogy.

		Actual Values	
		1	0
Predicted Values	1	TRUE POSITIVE 	FALSE POSITIVE 
	0	FALSE NEGATIVE 	TRUE NEGATIVE 

True Positive:

Interpretation: You predicted positive and it’s true.

You predicted that a woman is pregnant and she actually is.

True Negative:

Interpretation: You predicted negative and it’s true.

You predicted that a man is not pregnant and he actually is not.

False Positive: (Type 1 Error)

Interpretation: You predicted positive and it’s false.

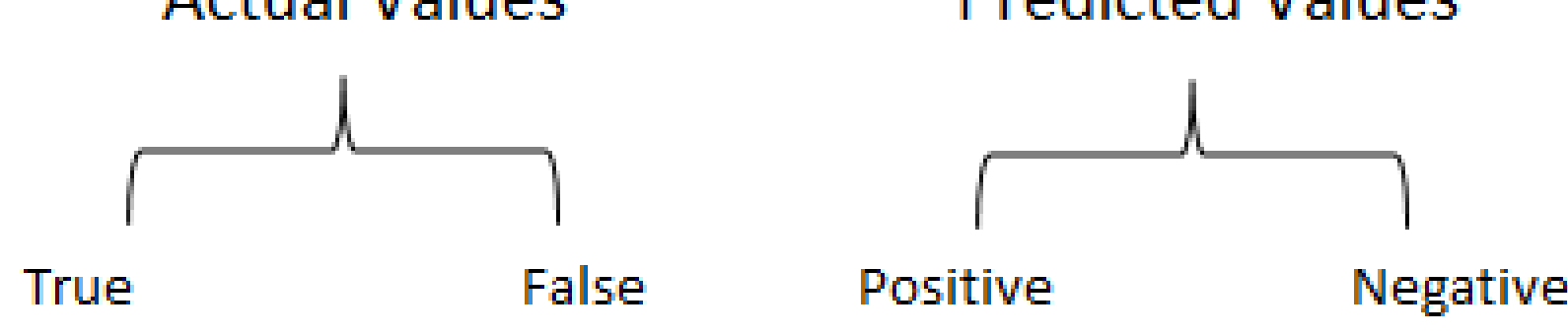
You predicted that a man is pregnant but he actually is not.

False Negative: (Type 2 Error)

Interpretation: You predicted negative and it’s false.

You predicted that a woman is not pregnant but she actually is.

Just Remember, We describe predicted values as Positive and Negative and actual values as True and False.



How to Calculate Confusion Matrix for a 2-class classification problem?

Let’s understand confusion matrix through math.



Recall



Out of all the positive classes, how much we predicted correctly. It should be high as possible.

Precision



Out of all the positive classes we have predicted correctly, how many are actually positive.

and **Accuracy** will be

Out of all the classes, how much we predicted correctly, which will be, in this case 4/7. It should be high as possible.

F-measure



It is difficult to compare two models with low precision and high recall or vice versa. So to make them comparable, we use F-Score. F-score helps to measure Recall and Precision at the same time. It uses Harmonic Mean in place of Arithmetic Mean by punishing the extreme values more.

...

I hope I’ve given you some basic understanding on what exactly is confusing matrix. If you like this post, a tad of extra motivation will be helpful by giving this post some claps 🙌. I am always open for your questions and suggestions. You can share this on Facebook, Twitter, Linkedin, so someone in need might stumble upon this.

You can reach me at:

Linked In : <https://www.linkedin.com/in/narkhedeSarang/>

Twitter : https://twitter.com/narkhede_sarang

Github : <https://github.com/TheSarang>

Thanks for Reading!

Machine LearningConfusion MartixRecallPrecisionData Science



5K claps

🐦🌐📘📌🔗



WRITTEN BY

Sarang Narkhede

Community Organizer @GDGRochester. Live and breath ML. All views are my own. Graduate CS student at @RIT.

Follow



Towards Data Science

A Medium publication sharing concepts, ideas, and codes.

Follow

See responses (45)

Discover Medium

Well come to a place where words matter. On Medium, smart voices and original ideas take center stage - with no ads in sight. [Watch](#)

Make Medium yours

Follow all the topics you care about, and we'll deliver the best stories for you to your homepage and inbox. [Explore](#)

Become a member

Get unlimited access to the best stories on Medium — and support writers while you're at it. Just \$5/month. [Upgrade](#)