✦ Member-only story

How to Learn the Definitions of Precision and Recall (For Good)

Why can I never remember the meaning of these?!

Adam Shafi · Follow

Published in Towards Data Science

6 min read · Jan 14, 2022

▶ Listen   ⬆ Share   ••• More

Photo by Oliver Buchmann on Unsplash

Introduction

Maybe you're new to Data Science or maybe, like me, you've got *some* Data Science experience and a **terrible** memory. Picture this... You're training models, running experiments and getting ready to make a deployment. You've got a product team

around you who have built a new app features and you want them to integrate your new classification model into the next build. You've shared some of your metrics. Then disaster strikes; *someone asks you to explain what precision and recall are...*

There are always terms and concepts that **I simply can't remember. Precision** and **Recall** are great examples. This article is mainly for me to finally learn these definitions, but hopefully you can benefit from my poor memory!

This article covers different ways to learn and communicate these concepts. To get the most out of this article, you should have some experience with training and scoring simple classification models.

Table of Contents

Start with the Confusion Matrix

Regardless of the method you eventually use to explain or remember **Precision** and **Recall,** an essential place to start is with the **Confusion Matrix.**

PREDICTED

	POSITIVE	NEGATIVE
ACTUAL POSITIVES	**TRUE** POSITIVES	**FALSE** NEGATIVES
ACTUAL NEGATIVE	**FALSE** POSITIVES	**TRUE** NEGATIVES

A standard confusion matrix. Image by Author

A reminder of the calculations

We can now write out our calculations in relation to the above matrix.

Precision

Precision is calculated by dividing the true positives by anything that was predicted as a positive.

$$\frac{\text{TRUE POSITIVES}}{\text{TRUE POSITIVES} + \text{FALSE POSITIVES}}$$

Recall

Recall (or True Positive Rate) is calculated by dividing the true positives by anything that should have been predicted as positive.

$$\frac{\text{TRUE POSITIVES}}{\text{TRUE POSITIVES} + \text{FALSE NEGATIVES}}$$

False Positive Rate

For completeness, let's also take a look at the False Positive Rate calculation.
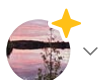
$$\frac{\text{FALSE POSITIVES}}{\text{FALSE POSITIVES} + \text{TRUE NEGATIVES}}$$
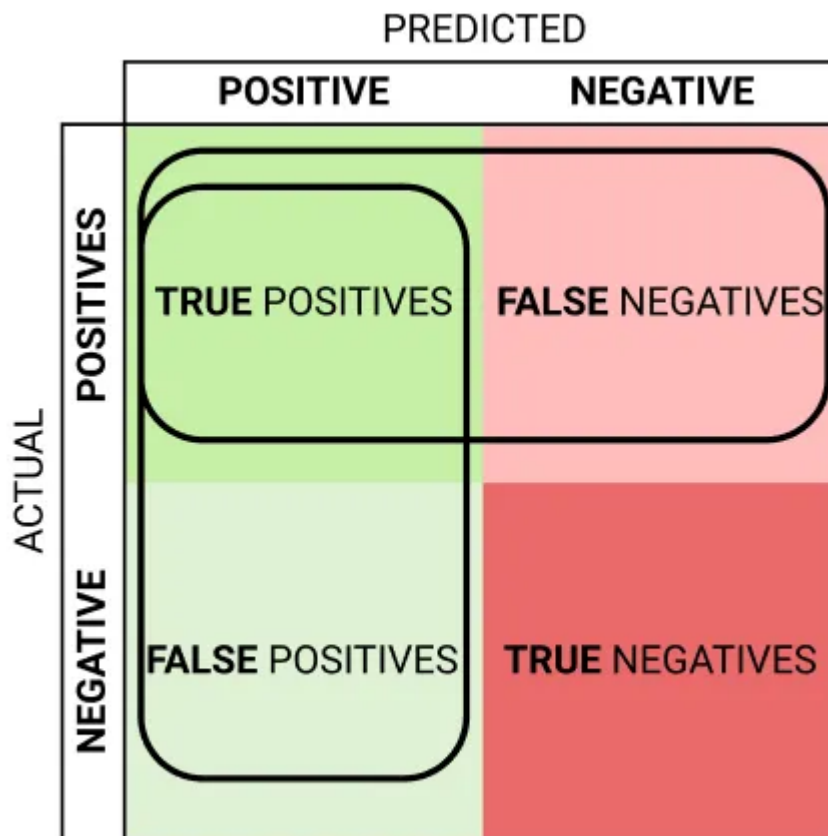
We can visually relate these back to the confusion matrix by circling the relevant fields.

Open in app ↗

Search Medium

We've now defined **Precision** and **Recall** and related these back to the confusion matrix. At this point I've explained the metrics and made a start on some visual ways to memorise these. Lets now take a look at clearly definitely these terms, which may help if you aren't a visual learner.

As a Sentence

Deepai.org have great definitions, which I've adapted and included in quotes below.

We can also frame these as questions which I find a particularly useful way to remember these.

Precision

> *The number of instances that are **relevant**, out of the **total instances the model retrieved.***

How many retrieved items are relevant?

Recall

> *The number of instances which the model **correctly identified as relevant** out of the total relevant instances.*

How many relevant items are retrieved?

An Analogy

When explaining concepts to non-technical users, analogies are always a useful approach. For some, this can be a great way to memorise something too. Reddit user u/question_23 used a common analogy for Precision and Recall in a very nice way in this post, which I've paraphrased below:

> *Explain it like fishing with a net. You use a wide net, and catch 80 of 100 total fish in a lake. That's 80% recall. But you also get 80 rocks in your net. That means 50% precision, half of the net's contents is junk.*
>
> *You could use a smaller net and target one pocket of the lake where there are lots of fish and no rocks, but you might only get 20 of the fish in order to get 0 rocks. That is 20% recall and 100% precision.*

Putting it into Context

Fully understanding a topic helps with memory, let's take a look at some common situations that relate to Precision and Recall.

A Calculation

Let's say you are trying to **predict customer churn**, using a classification model and some data. You've trained your model and made some predictions against a test dataset. These are the results:

n = 400		Predicted	
		Churn	Not Churn
Actual	Churn	300	35
	Not Churn	15	50

Results of a customer churn model. Image by Author.

Precision = 300/(300+15) = 95.2%

Recall = 300/(300+35) = 89.6%

What Does 100% Look Like?

Continuing on from above, what would 100% of both these metrics look like?

100% Precision: No false positives, every positive prediction is correct.

100% Recall: No false negatives, every negative prediction is correct.

Relationship to Other Common Metrics

F1 Score

The F1 score is simply the harmonic mean of Precision and Recall. We can also consider this in terms of true positives and negatives.

F1 Score formula. Image by Author.

ROC AUC

The ROC (receiver operating characteristic) curve is a great way to visualise the performance of a classification model and the area under the curve (AUC) is a very common metric used to compare different machine learning models.

To plot the ROC curve, we plot the True Positive Rate against the False Positive Rate for different classification probability thresholds (e.g. >50% = *True* vs >90%=*True*).

Remember, the True Positive Rate is the same as recall.

When to use these metrics?

This blog post by neptune.ai has a great discussion on when to use these different metrics.

https://neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc

A few more visuals

I've also scoured the internet for some other great ways to visualise these concepts. Hopefully these really hammer the point home.

Wikipedia has a good visualisation using circles, squares and colours (watch out for the difference between specificity and precision in this article):

Sensitivity and specificity — Wikipedia

Sensitivity and specificity mathematically describe the accuracy of a test which reports the presence or absence of a...

en.wikipedia.org

This Medium article also has some really nice Venn diagrams:

Explaining precision and recall

The first days and weeks of getting into NLP, I had a hard time grasping the concepts of precision, recall and...

medium.com

This is also a good example and explanation

Precision and Recall in Information Retrieval

I was chatting with a colleague recently. He worked in Bing Search. He mentioned precision and recall. This reminded me...

jamesmccaffrey.wordpress.com

Conclusion

The aim of this post was for me to finally memorise Precision and Recall, and honestly — I think it has helped. Take a look through these definitions, questions and visualisations and see what helps you solidify this knowledge in your brain.

Are there any machine learning concepts you can never remember, but should? Let me know in the comments!

Learn More

Clearing up Logistic Regression

The Guide for Non-Mathematicians

towardsdatascience.com

Random Forest for a non-Technical Audience

One of the most widely used algorithms today is actually quite tricky to explain...

towardsdatascience.com

Data Science   Machine Learning   Metrics   Data Analysis   Classification

tds

Follow

Written by Adam Shafi

483 Followers   ·   Writer for Towards Data Science

Data Scientist | Get in touch: https://www.linkedin.com/in/adamshafi

More from Adam Shafi and Towards Data Science

Adam Shafi **in** Towards Data Science

What is a Generalised Additive Model?

✦ · 8 min read · May 18, 2021

👏 251   💬 4   🔖⁺   •••

Jacob Marks, Ph.D. in Towards Data Science

How I Turned My Company's Docs into a Searchable Database with OpenAI

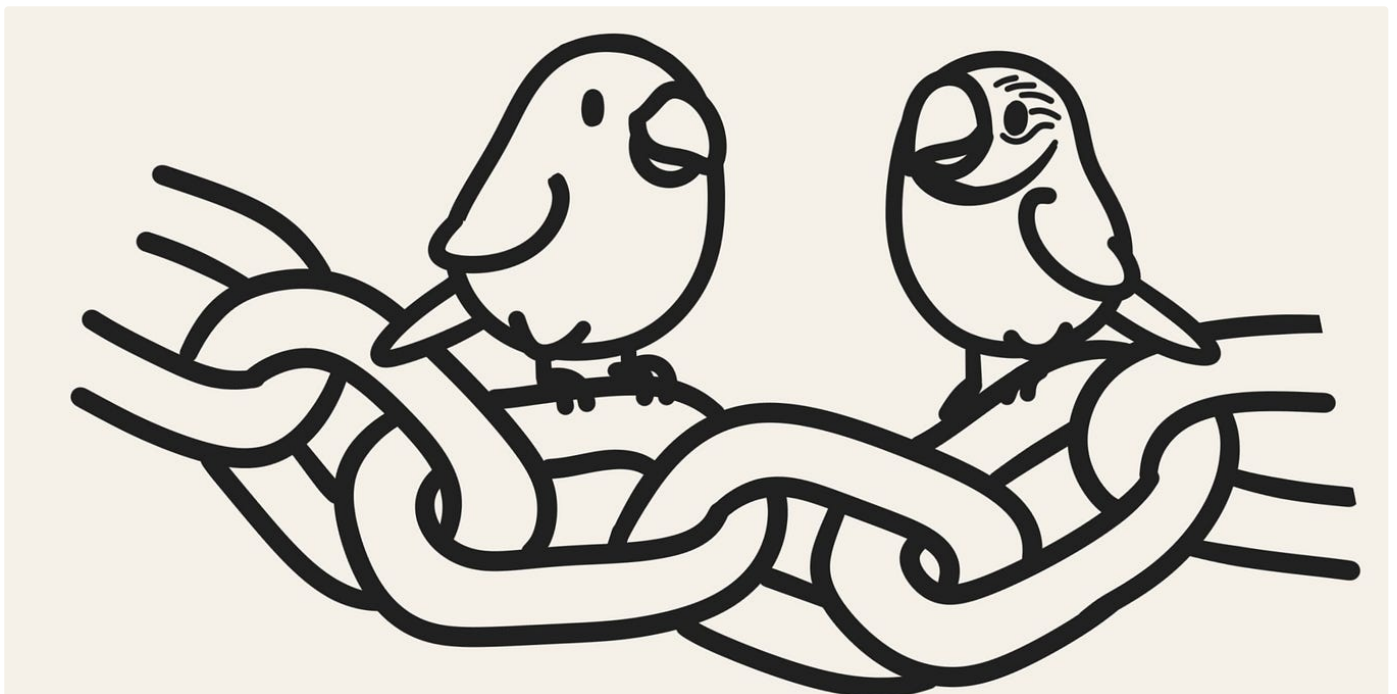And how you can do the same with your docs

15 min read · Apr 25

2.3K   30

Leonie Monigatti in Towards Data Science

Getting Started with LangChain: A Beginner's Guide to Building LLM-Powered Applications

A LangChain tutorial to build anything with large language models in Python

✦ · 12 min read · Apr 25

👏 1.3K   💬 13   🔖⁺   ···

Adam Shafi in Towards Data Science

Generating a Requirements.txt File from a Jupyter Notebook

A much overlooked way to save some time.

✦ · 3 min read · Jul 22, 2021

👏 306   💬 3   🔖⁺   ···

See all from Adam Shafi

See all from Towards Data Science

Recommended from Medium

Matt Chapman in Towards Data Science

The Portfolio that Got Me a Data Scientist Job

Spoiler alert: It was surprisingly easy (and free) to make

✦ · 10 min read · Mar 24

👏 2.8K   💬 44   🔖⁺   ⋯
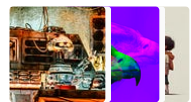
Albers Uzila in Level Up Coding

Wanna Break into Data Science in 2023? Think Twice!

It won't be smooth sailing for you

✦ · 11 min read · Dec 23, 2022

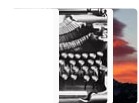👏 832   💬 14   🔖⁺   ⋯

Lists

What is ChatGPT?
9 stories · 28 saves

Staff Picks
303 stories · 65 saves

Stories to Help You Level-Up at Work
19 stories · 24 saves

Alex Mathers

14 things I wish I knew at 25 (now that I'm 38)

I'm writing soon after my 38th birthday. It's a time to reflect.

✦ · 4 min read · Dec 27, 2022

👏 11.9K   💬 212

Nikhil

BERT: Handling class imbalance in text classification

Most algorithms are designed to work best when the data is balanced, meaning that all classes are represented roughly equally. If there is...

3 min read · Dec 6, 2022

👏 61   💬 3

Serafeim Loukas, PhD in Towards AI

How To Estimate FP, FN, TP, TN, TPR, TNR, FPR, FNR & Accuracy for Multi-Class Data in Python in 5...

In this post, I explain how someone can read a confusion matrix and how to extract several performance metrics for a multi-class...

✨ · 6 min read · Feb 5

👏 64   💬

Leihua Ye, PhD

Why Data Scientists Should Learn Causal Inference

Climb up the ladder of causation

✦ · 7 min read · Jul 5, 2022

See more recommendations