Introduction to Probability Calibration:

Mansi Sarda · Following

2 min read · Jan 9, 2022

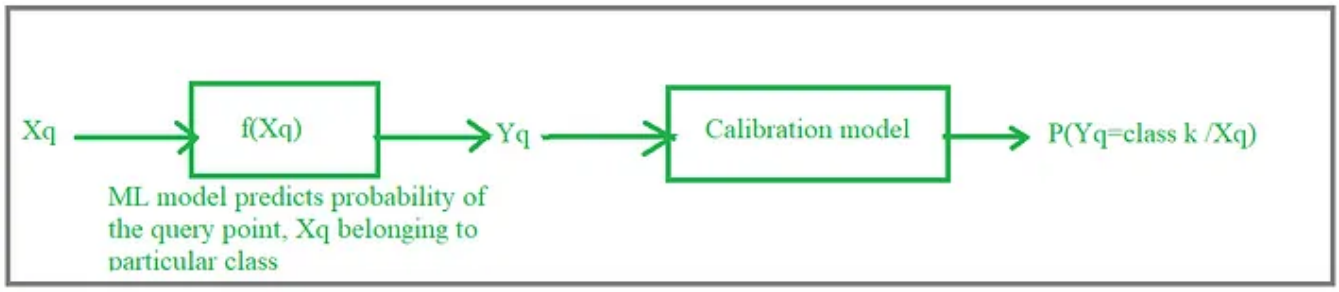
79

Probability Calibration refers to a method of calibrating probabilities which are predicted by various machine learning models like Gaussian Naive Bayes in order to have most accurate probabilities.

CalibratedClassifierCV outputs the calibrated probabilities using two methods — sigmoid and isotonic.

Why we need Probability Calibration?

Some important metrics such as logloss, Brier score takes the probability predicted by the model in order to find the loss. In such cases the probabilities which are not calibrated may not give good results. In addition to that not all classifiers provide well-calibrated probabilities, some being over-confident while others being under-confident. Thus, a separate calibration of predicted probabilities is desirable as a postprocessing.



When to apply Probability Calibration

Methods/Types of Probability Calibration provided by CalibratedClassifierCV :

1. Sigmoid — Sigmoid calibration is also called Platt Calibration.

$$P(y = 1|x) = \frac{1}{1 + \exp(Af(x) + B)}$$

As given in the above picture, f(x) is the predicted probability. P(y=1|x) is the calibrated probability. A and B are two scalar parameters that are learned by the algorithm which is calculated by solving an optimization problem. Sigmoid is only useful in cases where the actual probabilities are also sigmoidal in nature.

Cross Validation or cv parameter of CalibratedClassifierCV :

This determines the cross validation splitting strategy. It either takes any integer as cv in order to perform k-fold cross validation or “prefit” is passed

if the entire data is to be used for calibration.

Written by Mansi Sarda

Following

26 Followers

Pursuing MS in CS || Actively looking for software engineer roles

More from Mansi Sarda

Mansi Sarda in Analytics Vidhya

Scene Text Detection, Recognition and Translation.

Table of Contents :-

8 min read · Mar 21, 2021

115 3

Mansi Sarda in Analytics Vidhya

Quasar Detection using Machine Learning and Deep Learning Model

Table of Contents :-

13 min read · Aug 23, 2021

112

See all from Mansi Sarda

Recommended from Medium

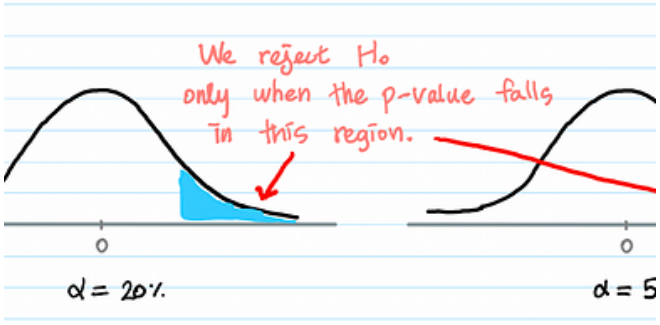
Michael At... in Stanford CS224W GraphML Tutor...

Tackling the Traveling Salesman Problem with Graph Neural...

Do you ever feel like there aren't enough hours in the day to get everything done? You...

12 min read · May 16

75



Ms Aerin in IntuitionMath

Chi Square Test—Intuition, Examples, and Step-by-Step...

The best way to see if two variables are related.

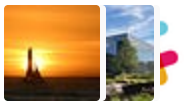
· 15 min read · Feb 13

407 3

Lists

Staff Picks

469 stories · 338 saves



Stories to Help You Level-Up at Work

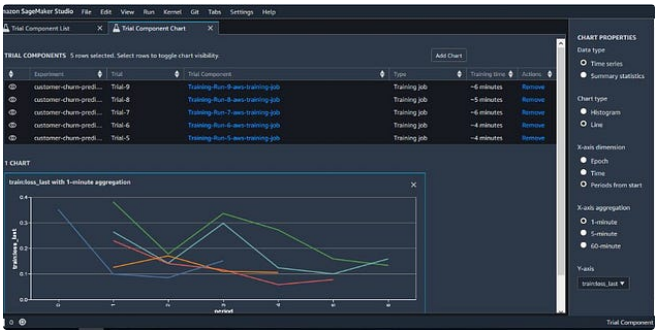
19 stories · 239 saves

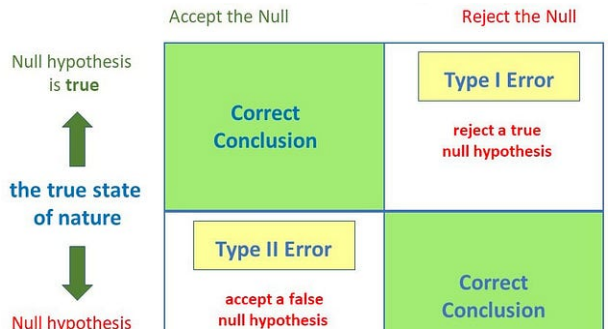
Self-Improvement 101

20 stories · 690 saves

Productivity 101

20 stories · 625 saves





Yugank Aman

Top MLOps Tools to Manage Machine Learning Lifecycle

Businesses continue transforming their operations to increase productivity and...

10 min read · May 30

94

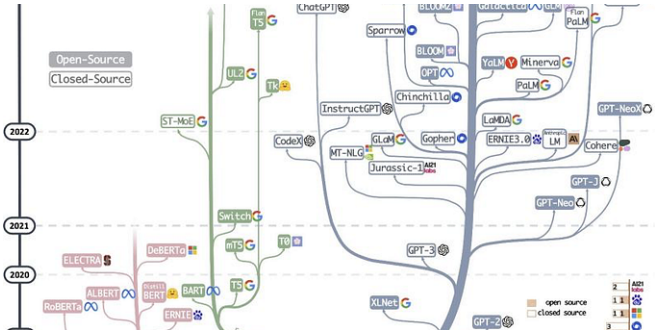
misun_song

Hypothesis Test and P-Value

Discerning Real Differences with P-values

7 min read · Sep 24

139



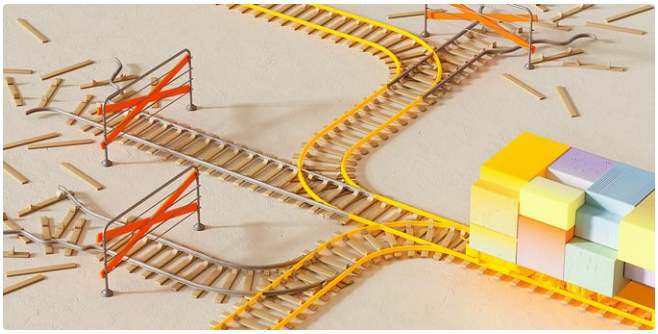
Haifeng Li

A Tutorial on LLM

Generative artificial intelligence (GenAI), especially ChatGPT, captures everyone's...

15 min read · Sep 14

805



Andrew Blance in Better Programming

MLOps and MLflops

How to understand MLOps architectures and diagrams

9 min read · Apr 3

496 3

See more recommendations