

Home

PUBLIC

Questions

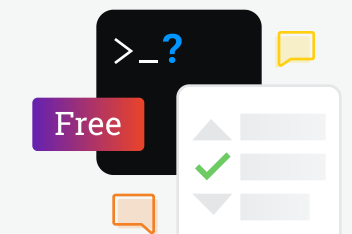
Tags

Users

Unanswered

TEAMS

Stack Overflow for Teams – Start collaborating and sharing organizational knowledge.



Create a free Team

Why Teams?

Why isn't Akaike information criterion used more in machine learning?

Asked 4 years, 10 months ago Modified 3 years ago Viewed 6k times

▲

24

▼

7

machine-learning model-selection aic bic

Share Cite Improve this question

Follow

edited Nov 16, 2017 at 6:29

asked Nov 15, 2017 at 14:12

Richard Hardy

57.3k

11

107

225

echo

921

7

13

10

because nobody's calculating the likelihoods? – Aksakal Nov 15, 2017 at 14:16

1

What do you mean by "contemporary machine learning methods"? As far as I used AIC and BIC are used frequently. – Ferdi Nov 15, 2017 at 14:16

5

Also why the -1? Remember there are no stupid questions -- each question tries to shed light on the universe – echo Nov 15, 2017 at 14:28

4

@echo: I didn't downvote, but I think your question would be improved if you could source/support the main claim (that machine learning methods do take advantage of these BIC and AIC model selection criteria) – user603 Nov 15, 2017 at 14:32

2

@Aksakal Thanks. I think it is better if questions constructed around a sweeping claim could source that claim. I mean as a general rule. – user603 Nov 15, 2017 at 14:46

Show 4 more comments

1 Answer

Sorted by: Highest score (default)

▲

20

▼

Share Cite Improve this answer

Follow

edited Sep 11, 2019 at 9:31

answered Jan 26, 2018 at 20:58

LiKao

2,349

1

18

25

Care to chime in on this post stats.stackexchange.com/questions/325129/... ? I've haven't had any luck with it for a while. – Skander H. Jan 26, 2018 at 21:42

1

@LiKao Can you cite references on the "techniques" of hiding parameters, like the case of intersecting digits. – horaceT Sep 1, 2019 at 22:16

@horaceT Unfortunately I don't know of any paper, that gives this example. In papers on MDL there is the notion of "functional complexity" (e.g. lpl.psy.ohio-state.edu/documents/MNP.pdf see eq 10). Often the example is made with constrained parameters (e.g. researchgate.net/publication/...). I like to turn the example around when discussing this, and show that a complex single parameter can capture multiple simple parameters because I find it more intuitive. – LiKao Sep 9, 2019 at 12:04

@horaceT Also, if you like a more mathematical treatment, consider that it is proven that filling curves exist, i.e. there is a bijection $f_{1,2} : \mathbb{R} \rightarrow \mathbb{R}^2$. This bijection can easily be used to define bijections $f_{1,N} : \mathbb{R} \rightarrow \mathbb{R}^N$. So for any model with N parameters, I can use $f_{1,N}$ to first get an N dimensional vector from my single parameter, then supply this vector as a parameter to the N parameter model. This gives me a functional equivalent 1 parameter model. Fitting that model, however would be very complicated at least. – LiKao Sep 9, 2019 at 12:11

@LiKao This' quite fascinating. Pls reference said proof of "filing curves". I could see that constrained parameters has "less" degree of freedom. Naively, if $f(x,y) = 0$, y is just a function of x ; you just put $g(x)$ where y is. Can't you do similar things with constrained optimization. – horaceT Sep 9, 2019 at 21:37

Show 11 more comments

Your Answer

B I

Sign up or log in

Sign up using Google

Sign up using Facebook

Sign up using Email and Password

Post as a guest

Name

Email

Required, but never shown

Post Your Answer By clicking "Post Your Answer", you agree to our [terms of service](#), [privacy policy](#) and [cookie policy](#)

Not the answer you're looking for? Browse other questions tagged [machine-learning](#) [model-selection](#) [aic](#) [bic](#) or [ask your own question](#).

Featured on Meta

☐

Announcing the Stack Overflow Student Ambassador Program

☐

Google Analytics 4 (GA4) upgrade



Linked

- 13

Is it possible that two Random Variables from the same distribution family have the same expectation and variance, but different higher moments?
- 5

Comparison between MDL and BIC
- 9

How can the AIC or BIC be used instead of the train/test split?
- 3

Constructive mathematics and statistics
- 2

What to do if the rate of misclassification in Bayesian model selection depends on the model parameter

Related

- 23

AIC & BIC number interpretation
- 6

In Bayesian Information Criterion (BIC), why does having bigger n get penalized?
- 1

Where is the divide between information criterion (AIC, BIC, etc...) and cross validation?
- 7

Why is AIC or BIC commonly used in model selections for time series forecasting?
- 19

Why does the Akaike Information Criterion (AIC) sometimes favor an overfitted model?
- 3

Model selection criteria that represent a compromise between AIC and BIC

Hot Network Questions

- ask

Compiling C++20 program that uses modules with Ubuntu 22.04
- I closed my CU checking/savings accounts. Weeks later the bank auto-paid my rent. Now they want me to pay them back
- Usage of "the" before "truth"
- A Logician, three students and three initials
- Why don't bridge players try to surprise their opponents by learning a different bidding system?
- [S]

During the Iran-Iraq war, did the Iranians use small children to trip landmines planted by Iraq?
- Stunted evolution on a tidally locked planet around a Red Dwarf?
- Short story, probably from the '70s, where a new procedure would supposedly make people smarter
- Which of those parameters is default in DBCC SHRINKFILE command?
- Why is the ground warm during winter and vice-versa?
- Staying more than 90 days in Schengen area as a French citizen, but only have a US passport
- conceptual understanding of quadratic regression
- Pack of dogs - best response to owner?
- Why underpromote? Hjartarson vs Andersen
- How many undiscovered planets might be in our solar system?
- Being TA for abusive professor
- How do airplanes with thrust lines above CG still fly without a down pitch moment?
- Is it necessary to have a variety of publishers in the list of my own publications?
- Novel with time travel, religious references, main character being called "Gog," aliens, and pre-missing-link humanoids. Likely from 1980s
- Lady Penelope's birth date and wikipedia
- What value is there in requiring students to answer word problems in complete sentences?
- Get Official Microsoft Office Professional Plus 2010
- Is there a duration or rest error in the last bar here?
- A question about DR/BDR routers in OSPF

[Question feed](#)

CROSS VALIDATED

- Tour
- Help
- Chat
- Contact
- Feedback

COMPANY

- Stack Overflow
- Teams
- Advertising
- Collectives
- Talent
- About
- Press
- Legal
- Privacy Policy
- Terms of Service
- Cookie Settings
- Cookie Policy

STACK EXCHANGE NETWORK

- Technology
- Culture & recreation
- Life & arts
- Science
- Professional
- Business
- API
- Data

[Blog](#) [Facebook](#) [Twitter](#) [LinkedIn](#) [Instagram](#)