Google Discussions

EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 157 DISCUSSION

Actual exam question from Google's Professional Data Engineer
Question #: 157
Topic #: 1
[All Professional Data Engineer Questions]

Your team is working on a binary classification problem. You have trained a support vector machine (SVM) classifier with default parameters, and received an area under the Curve (AUC) of 0.87 on the validation set. You want to increase the AUC of the model. What should you do?

A. Perform hyperparameter tuning

B. Train a classifier with deep neural networks, because neural networks would always beat SVMs

C. Deploy the model and measure the real-world AUC; it's always higher because of generalization

D. Scale predictions you get out of the model (tune a scaling factor as a hyperparameter) in order to get the highest AUC

Show Suggested Answer

by [deleted] at *March 22, 2020, 7:33 a.m.*

Comments

Type your comment...

Submit

aadaisme Highly Voted 4 years, 4 months ago
Seems to be A. Preprocessing/scaling should be done with input features, instead of predictions (output)
upvoted 42 times

upvoted 42 times

FARR Highly Voted 4 years, 2 months ago

A
Deep LEarning is not always the best solution
D talks about fudgin the output which is wrong

upvoted 11 times

MaxNRG Most Recent 10 months, 2 weeks ago

Selected Answer: A

https://www.quora.com/How-can-I-improve-Precision-Recall-AUC-under-Imbalanced-Classification

upvoted 3 times

vaga1 1 year, 5 months ago

Selected Answer: A

B,C are simply not true. D is modifing the scoring, making it not realiable anymore. A makes sense, is potentially increasing the model accuracy.

upvoted 2 times

rishu2 1 year, 5 months ago

Selected Answer: A

a is the correct answer

upvoted 1 times

musumusu 1 year, 8 months ago

Answer A,
why not B, Deep Neu Net. are better for sure but AUC is 0.87 is already good. Don't go for complex and time taking model. If AUC more than 0.95, it can be a reason of overfit.
Now just check SVM params for hypertuning if you can bring it close to 0,9-0,95

upvoted 1 times

Kvk117 1 year, 9 months ago

Selected Answer: A

a is the correct answer

upvoted 1 times

Dan137 1 year, 12 months ago

Also a good read is: https://cloud.google.com/ai-platform/training/docs/hyperparameter-tuning-overview

upvoted 1 times

medeis_jar 2 years, 10 months ago

Selected Answer: A

as mentioned by Spider7 "performing tuning rather than using the model default parameters there's a way to increase the overall model performance --> A."

upvoted 2 times

JG123 2 years, 11 months ago

Correct: A

upvoted 1 times

Spider7 2 years, 11 months ago

0.89 it's already not bad but by performing tuning rather then using the model default parameters there's a way to increase the overall model performance --> A.

upvoted 3 times

Spider7 2 years, 11 months ago

0.87 precisely

upvoted 1 times

hdmi_switch 3 years, 3 months ago

Not C because real-world AUC value falls between 0.5 and 1.0 usually, this wouldn't help.

A seems the most straigh forward.

upvoted 3 times

Mitra123 3 years, 7 months ago

For a large enough training set DNN will most likely beat a SVM. However the opposite may or may not be true. It also depends on the complexity of the problem. Which we don't know from the question. For image, nlp, I say B can be a good answer
However, if we decide to stick with SVM, D reduces overfitting and may increase AUC.
I am torn between the two!

upvoted 1 times

ArunSingh1028 3 years, 8 months ago

Ans - D when the model is overfitted means want to increase the AUC, we always perform hyperparameter tuning, Increase regularisations, decrease input feature parameters etc.

upvoted 1 times

nitinbhatia 4 years ago

AUC is scale-invariant. It measures how well predictions are ranked, rather than their absolute values. So answer shall be A
https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=en

upvoted 2 times

arghya13 4 years ago

Definitely not D
https://developers.google.com/machine-learning/crash-course/classification/check-your-understanding-roc-and-auc

upvoted 3 times

saurabh1805 4 years, 2 months ago

A for me, read below link for more details.

https://towardsdatascience.com/understanding-hyperparameters-and-its-optimisation-techniques-f0debba07568

upvoted 4 times

Load full discussion...