

Google Discussions



Exam Professional Machine Learning Engineer All Questions

View all questions & answers for the Professional Machine Learning Engineer exam

Go to Exam

EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 3 DISCUSSION

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 3

Topic #: 1

[\[All Professional Machine Learning Engineer Questions\]](#)

You were asked to investigate failures of a production line component based on sensor readings. After receiving the dataset, you discover that less than 1% of the readings are positive examples representing failure incidents. You have tried to train several classification models, but none of them converge. How should you resolve the class imbalance problem?

- A. Use the class distribution to generate 10% positive examples.
- B. Use a convolutional neural network with max pooling and softmax activation.
- C. Downsample the data with upweighting to create a sample with 10% positive examples.
- D. Remove negative examples until the numbers of positive and negative examples are equal.

Show Suggested Answer

by [inder0007](#) at June 2, 2021, 9:48 p.m.

Comments

Type your comment...

Submit

[celia20200410](#) **Highly Voted** 3 years, 3 months ago

ANS: C

<https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data#downsampling-and-upweighting>

- less than 1% of the readings are positive
- none of them converge.

Downsampling (in this context) means training on a disproportionately low subset of the majority class examples.

   upvoted 31 times

  **mousseUwU** 3 years ago

Agree, C is correct

   upvoted 2 times

  **MisterHairy** Highly Voted  3 weeks, 6 days ago

=New Question3=

You are going to train a DNN regression model with Keras APJs using this code:

```
model = tf.keras.Sequential()
model.add(tf.keras.layers.Dense(
    256,
    use_bias=True,
    activation='relu',
    kernel_initializer=None,
    kernel_regularizer=None,
    input_shape=(500,)))
model.add(tf.keras.layers.Dropout(rate=0.25))
model.add(tf.keras.layers.Dense(
    128, use_bias=True,
    activation='relu',
    kernel_initializer='uniform',
    kernel_regularizer='l2'))
model.add(tf.keras.layers.Dropout(rate=0.25))
model.add(tf.keras.layers.Dense(
    2, use_bias=False,
    activation='softmax'))
model.compile(loss='mse')
```



How many trainable weights does your model have? (The arithmetic below is correct.)

- A. $501 \times 256 + 257 \times 128 + 2 = 161154$
- B. $500 \times 256 + 256 \times 128 + 128 \times 2 = 161024$
- C. $501 \times 256 + 257 \times 128 + 128 \times 2 = 161408$
- D. $500 \times 256 \times 0(?) + 25 \times 256 \times 128 \times 0(?) + 25 \times 128 \times 2 = 4044$

   upvoted 9 times

  **toooooony55** 2 years, 9 months ago

B: Dense layers with 100 % trainable weights, the dropout rate at 0.25 will randomly drop 25 % for the regularization's sake - still training for 100 % of the weights.

   upvoted 1 times

  **AlexZot** 2 years, 9 months ago

Correct answer is C. Do not forget about bias term which is also trainable parameter.

   upvoted 5 times

  **sakura65** 2 years, 6 months ago

Why 128 for the last layer is correct and not 129 X 2?

   upvoted 1 times

  **suresh_vn** 2 years, 2 months ago

because of use_bias = False

   upvoted 1 times




  **suresh_vn** 2 years, 2 months ago

C is correct. 2nd Layer with use_bias = True



   upvoted 3 times

  **NickHapton** 2 years, 10 months ago

Why do you post new questions in every existing question rather than post them as a new question?

   upvoted 4 times

  **MisterHairy** 2 years, 6 months ago

  **MisterHairy** 2 years, 9 months ago

Only moderator can post new questions. Thus, I am left with this format. I have emailed the additional questions to the moderator, but he/she has not added them to the site. These questions were received off of other practice tests, but answers were not provided.

   upvoted 4 times

  **MisterHairy** 2 years, 10 months ago

Answer?

   upvoted 1 times

  **Mohamed_Mossad** 2 years, 4 months ago

D , is the only option that takes care of the dropout factor

   upvoted 1 times

  **Mohamed_Mossad** 2 years, 4 months ago

my bad , this was tricky "The Dropout Layer randomly disables neurons during training. They still are present in your model and therefore aren't discounted from the number of parameters in your model summary." , so D is wrong , C and A takes care of the bias , but C is correct

   upvoted 2 times

[Load full discussion...](#)

  **ramen_lover** Most Recent 3 weeks, 6 days ago

The answer is C.

We have to note that

- (1) Downsampling on major class
 - (2) Upsampling on minor class
 - (3) Upweighting on minor class
- all work for imbalanced data.

However, the key assumption in the question is that "You have tried to train several classification models, but none of them converge".

You are not asked to tackle imbalanced data but asked to handle the non-convergence problem (due to the limited resources or the pooriness of the algorithm).

In the official document, it says: "If you have an imbalanced data set, first try training on the true distribution. If the model works well and generalizes, you're done! If not, try the following downsampling and upweighting technique."

<https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data#downsampling-and-upweighting>

In other words, the "downsampling and upweighting technique" is the technique for the non-convergence problem (not for for the imbalanced data).

   upvoted 5 times

  **Fatiy** 3 weeks, 6 days ago

Selected Answer: C

C. Downsample the data with upweighting to create a sample with 10% positive examples.

Dealing with class imbalance can be challenging for machine learning models. One common approach to resolving the problem is to downsample the data, either by removing examples from the majority class or by oversampling the minority class. In this case, since you have very few positive examples, you would want to oversample the positive examples to create a sample that better represents the underlying distribution of the data. This could involve using upweighting, where positive examples are given a higher weight in the loss function to compensate for their relative scarcity in the data. This can help the model to better focus on the positive examples and improve its performance in classifying failure incidents.

   upvoted 2 times

  **PhilipKoku** 4 months, 2 weeks ago

Selected Answer: C

This approach involves downsampling the majority class (negative examples) and upweighting the minority class (positive examples) to create a balanced dataset.

By doing so, the model can learn from both classes effectively.


Reference: How to Handle Imbalanced Classes in Machine Learning [<https://elitedatascience.com/imbalanced-classes>]

   upvoted 1 times

  **fragkris** 10 months, 3 weeks ago

Selected Answer: C

C - Downsample the majority and add weights to it.


   upvoted 2 times

  **tatpicc** 11 months, 1 week ago

Max Pooling is a pooling operation that calculates the maximum value for patches of a feature map, and uses it to create a downsampled (pooled) feature map. It is usually used after a convolutional layer

downsampled (pooled) feature map. It is usually used after a convolutional layer.

   upvoted 1 times

  **M25** 1 year, 5 months ago

Selected Answer: C

Went with C

   upvoted 1 times

  **Puneet2022** 1 year, 5 months ago

Selected Answer: C

<https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data#downsampling-and-upweighting>



   upvoted 1 times

  **enghabeth** 1 year, 8 months ago

Selected Answer: C

<https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data>

   upvoted 1 times

  **Fatiy** 1 year, 8 months ago

C. Downsample the data with upweighting to create a sample with 10% positive examples.

Dealing with class imbalance can be challenging for machine learning models. One common approach to resolving the problem is to downsample the data, either by removing examples from the majority class or by oversampling the minority class. In this case, since you have very few positive examples, you would want to oversample the positive examples to create a sample that better represents the underlying distribution of the data. This could involve using upweighting, where positive examples are given a higher weight in the loss function to compensate for their relative scarcity in the data. This can help the model to better focus on the positive examples and improve its performance in classifying failure incidents.

   upvoted 1 times

  **SharathSH** 1 year, 9 months ago

Answer would obviously be C

As the dataset is imbalanced and you need to resolve this issue in order to obtain desired result the best approach will be to downsample the data.

   upvoted 1 times

  **EFIGO** 1 year, 11 months ago

Selected Answer: C

Best practice for imbalanced dataset is to downsample with upweight

<https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data#downsampling-and-upweighting>

   upvoted 1 times

  **GCP72** 2 years, 2 months ago

Selected Answer: C

Correct answer is "C"

   upvoted 1 times

  **enghabeth** 2 years, 2 months ago

Selected Answer: C

C. because regardless of the model you use, you should always try to transform or adapt your dataset so that it is more balanced

   upvoted 1 times

  **Mohamed_Mossad** 2 years, 4 months ago

Selected Answer: C



<https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data>

   upvoted 1 times

  **Mohamed_Mossad** 2 years, 4 months ago

Selected Answer: D

<https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data>

   upvoted 1 times

  **Mohamed_Mossad** 2 years, 4 months ago

sorry , mean C

   upvoted 1 times

[Load full discussion...](#)

Start Learning for free



Social Media

[Facebook](#) , [Twitter](#)

[YouTube](#) , [Reddit](#)

[Pinterest](#)



We are the biggest and most updated IT certification exam material website.

Using our own resources, we strive to strengthen the IT professionals community for free.



© 2024 ExamTopics

ExamTopics doesn't offer Real Microsoft Exam Questions. ExamTopics doesn't offer Real Amazon Exam Questions. ExamTopics Materials do not contain actual questions and answers from Cisco's Certification Exams.

CFA Institute does not endorse, promote or warrant the accuracy or quality of ExamTopics. CFA® and Chartered Financial Analyst® are registered trademarks owned by CFA Institute.

