

Questionnaire (5th chapter)

Total de pontos 0/0 ?

O e-mail do participante (**breno.xavier@ufpe.br**) foi registrado durante o envio deste formulário.

Why do we first resize to a large size on the CPU, and then to a smaller size on the GPU? *

Resizing images is a crucial preprocessing step in deep learning. It's common to first resize the images to a larger size on the CPU, as it's faster and less memory-intensive. Then, the resized images are further reduced in size on the GPU, which allows the data to fit into memory and speeds up the training process

What are the two ways in which data is most commonly provided, for most deep learning datasets? *

Data is commonly provided in most deep learning datasets either as a folder of labeled images or as a CSV file with labels and file paths

Give two examples of ways that image transformations can degrade the quality of the data. *

Two examples of ways that image transformations can degrade the quality of the data are: first: rotation can cause loss of information or distortion of features, second: esizing an image to a smaller size can result in loss of detail and important information

What method does fastai provide to view the data in a DataLoaders? *

show_batch

What method does fastai provide to help you debug a DataBlock? *

summary



Should you hold off on training a model until you have thoroughly cleaned your data? *

yes, i should

What are the two pieces that are combined into cross-entropy loss in PyTorch? *

The two pieces that are combined into cross-entropy loss in PyTorch are negative log likelihood loss and log softmax activation function

What are the two properties of activations that softmax ensures? Why is this important? *

The two properties of activations that softmax ensures are that they are non-negative and sum to 1, it is important because it allows us to interpret the activations as probabilities and use them to make predictions

When might you want your activations to not have these two properties? *

We might want activations without these properties when dealing with regression or when we have prior knowledge of the data distribution

Why can't we use torch.where to create a loss function for datasets where our label can have more than two categories? *

imitation of torch.where in handling multi-class label makes it unsuitable for creating loss function for datasets with more than two categories

What is the value of $\log(-2)$? Why? *

it's "not a number" in real number system, The natural logarithm function is only defined for $x > 0$, not for negative values or zero, so the value of $\log(-2)$ is undefined or "NaN"



What are two good rules of thumb for picking a learning rate from the learning rate finder? *

Two good rules of thumb for picking a learning rate from the learning rate finder are selecting the point where the loss starts to decrease quickly and choosing a value approximately one order of magnitude less than the minimum loss

What two steps does the `fine_tune` method do? *

The `fine_tune` method in `fastai` combines freezing and training of model groups to fine-tune a pre-trained model on a new dataset

In Jupyter Notebook, how do you get the source code for a method or function? *

can get the source code for a method or function by appending a question mark to the end of the method/function name and running the cell

What are discriminative learning rates? *

Discriminative learning rates refer to the practice of using different learning rates for different layers of a neural network. This technique involves setting a higher learning rate for earlier layers (which learn low-level features) and a lower learning rate for later layers (which learn higher-level features), thereby improving the model's performance

How is a Python slice object interpreted when passed as a learning rate to `fastai`? *

When a slice object is passed as a learning rate to `fastai`, it is interpreted as a range of learning rates to use during training, with the first value indicating the lower bound and the second value indicating the upper bound. The learning rate is then increased linearly from the lower bound to the upper bound over the course of the training



Why is early stopping a poor choice when using 1 cycle training? *

Because 1 cycle training already has built-in annealing and reducing the learning rate early can cause the model to miss out on the high learning rate phase, which is critical for generalization

What is the difference between resnet50 and resnet101? *

The main difference between resnet50 and resnet101 is the number of layers, with resnet101 having a deeper architecture with more layers than resnet50

What does to_fp16 do? *

'to_fp16' is a method that converts model weights and activations to half-precision floating-point format in order to reduce memory usage and speed up training

Este formulário foi criado em Universidade Federal de Pernambuco.

Google Formulários

