- **Introduction to Programming Frameworks**
- **Lecture Notes (Optional)**
- **Quiz**
  - 테스트: Hyperparameter tuning, Batch Normalization, Programming Frameworks
    10개의 질문
- **Programming Assignment**
- **References & Acknowledgments**

테스트테스트 • 20 min20 minutes

Hyperparameter tuning, Batch Normalization, Programming Frameworks

과제 제출
기한년 9월 6일 오후 3:59 KST년 9월 6일 오후 3:59 KST
시도하기8 hours당 3회

다시 시도해주십시오

성적 받기
통과 점수:80% 이상
성적
100%

피드백 보기

최고 점수가 유지됩니다.

탐색 확인

이 페이지에서 나가시겠습니까?

이 페이지에 머물기 이 페이지에서 나가기

Hyperparameter tuning, Batch Normalization, Programming Frameworks
성적 평가 퀴즈 • 20 min

만료 년 9월 6일 오후 3:59 KST

축하합니다! 통과하셨습니다!
통과 점수: 80% 이상

학습 계속하기

성적
100%

Hyperparameter tuning, Batch Normalization, Programming Frameworks

최신 제출물 성적
100%
1.
질문 1

If searching among a large number of hyperparameters, you should try values in a grid rather than random values, so that you can carry out the search more systematically and not rely on chance. True or False?

1 / 1점

○ True

◉ False

✓
맞습니다
2.
질문 2

Every hyperparameter, if set poorly, can have a huge negative impact on training, and so all hyperparameters are about equally important to tune well. True or False?

1 / 1점
○○

True

◉◉

False

✓
맞습니다

Yes. We've seen in lecture that some hyperparameters, such as the learning rate, are more critical than others.

3.
질문 3

During hyperparameter search, whether you try to babysit one model (“Panda” strategy) or train a lot of models in parallel (“Caviar”) is largely determined by:

1 / 1점
◉◉

The amount of computational power you can access

○○

The number of hyperparameters you have to tune

○○

The presence of local minima (and saddle points) in your neural network

○○

Whether you use batch or mini-batch optimization

✓
맞습니다
4.
질문 4

If you think ββ (hyperparameter for momentum) is between 0.9 and 0.99, which of the following is the recommended way to sample a value for beta?

1 / 1점
○○

r = np.random.rand() beta = r*0.09 + 0.9

○○

r = np.random.rand() beta = 1-10**(- r + 1)

○○

r = np.random.rand() beta = r*0.9 + 0.09

◉◉

r = np.random.rand() beta = 1-10**(- r - 1)

✓
맞습니다
5.
질문 5

Finding good hyperparameter values is very time-consuming. So typically you should do it once at the start of the project, and try to find very good hyperparameters so that you don’t ever have to revisit tuning them again. True or false?

1 / 1점
◉◉

False

○○

True

✓
맞습니다
6.
질문 6

In batch normalization as presented in the videos, if you apply it on the $l$lth layer of your neural network, what are you normalizing?

1 / 1점
◉◉

$z^{[l]}$z[l]

○○

$a^{[l]}$a[l]

○○

$b^{[l]}$b[l]

○○

$W^{[l]}$W[l]

✓
맞습니다
7.
질문 7

In the normalization formula $z^{(i)}_{norm} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$znorm(i) = σ2+ε
√
z(i)−μ, why do we use epsilon?

1 / 1점

To have a more accurate normalization

To speed up convergence

To avoid division by zero

In case μ μ is too small

맞습니다
8.
질문 8

Which of the following statements about γ γ and β β in Batch Norm are true?

1 / 1점

They can be learned using Adam, Gradient descent with momentum, or RMSprop, not just with gradient descent.

맞습니다

There is one global value of $\gamma \in \Re$ γ ∈ ℜ and one global value of $\beta \in \Re$ β ∈ ℜ for each layer, and applies to all the hidden units in that layer.

β β and γ γ are hyperparameters of the algorithm, which we tune via random sampling.

They set the mean and variance of the linear variable $z^{[l]}$ z[l] of a given layer.

맞습니다

The optimal values are $\gamma = \sqrt{\sigma^2 + \varepsilon}$ γ = σ2 + ε
√
, and $\beta = \mu$ β = μ.

9.
질문 9

After training a neural network with Batch Norm, at test time, to evaluate the neural network on a new example you should:

1 / 1점

If you implemented Batch Norm on mini-batches of (say) 256 examples, then to evaluate on one test example, duplicate that example 256 times so that you're working with a mini-batch the same size as during training.

Skip the step where you normalize using μ μ and σ^2 σ2 since a single test example cannot be normalized.

Use the most recent mini-batch's value of μ μ and σ^2 σ2 to perform the needed normalizations.

Perform the needed normalizations, use μ μ and σ^2 σ2 estimated using an exponentially weighted average across mini-batches seen during training.

맞습니다
10.
질문 10

Which of these statements about deep learning programming frameworks are true? (Check all that apply)

1 / 1점

A programming framework allows you to code up deep learning algorithms with typically fewer lines of code than a lower-level language such as Python.

맞습니다

Even if a project is currently open source, good governance of the project helps ensure that the it remains open even in the long term, rather than become closed or modified to benefit only one company.

맞습니다

Deep learning programming frameworks require cloud-based machines to run.