

Machine Learning Crash Course – Week 3 Quiz Q&A

Q1- Suppose you are working on a spam classifier, where spam emails are positive examples ($y=1$) and non-spam emails are negative examples ($y=0$). You have a training set of emails in which 99% of the emails are non-spam and the other 1% is spam. Which of the following statements are true? Check all that apply.

- I- If you always predict spam (output $y=1$), your classifier will have a recall of 0% and precision of 99%.
- II- If you always predict non-spam (output $y=0$), your classifier will have an accuracy of 99%.
- III- If you always predict non-spam (output $y=0$), your classifier will have a recall of 0%.
- IV- If you always predict spam (output $y=1$), your classifier will have a recall of 100% and precision of 1%.
- V- A good classifier should have both a high precision and high recall on the cross-validation set.
- VI- If you always predict non-spam (output $y=0$), your classifier will have 99% accuracy on the training set, and it will likely perform similarly on the cross-validation set.
- VII- If you always predict non-spam (output $y=0$), your classifier will have 99% accuracy on the training set, but it will do much worse on the cross validation set because it has overfit the training data.

- I, II, III, V, VI
- II, III, IV, V, VII
- I, II, III, V, VII
- II, III, IV, V, VI ✓

A1-

- I- False - Her email spam dersek recall değerimiz 0% değil 100% olurken, precision değerimiz 99% değil %1 olur.
- II- True - Geçerli bir önermedir.
- III- True - Geçerli bir önermedir.
- IV- True - Geçerli bir önermedir.
- V- True - Geçerli bir önermedir.
- VI- True - Geçerli bir önermedir.
- VII- False - Her email spam değil dersek, train datası üzerinde accuracy değerimiz 99% olur, fakat cross validation set üzerinde yine aynı sonuç beklenir. Farklı bir sonuç beklenmez.

Q2- Which of the following statement is true?

- Without activation functions neural networks are linear ✓
- Adding hidden layers makes neural networks non-linear
- Using any mathematical function as an activation function makes neural networks non-linear
- ReLU activation function converts the weighted sum to a value between 0 and 1.
- Different layers should have only one type activation function.

A2- Aktivasyon fonksiyonları olmadan modeller $w \cdot x + b$ türünden olacağı için NN'ler lineer model olur. Gizli katman eklemek modelleri lineer yapmaz. Sadece non-lineer fonksiyonların aktivasyon fonksiyonu olarak kullanılması NN'leri non-lineer model yapar. Örneğin $1+x$ gibi bir aktivasyon fonksiyonu kullanmak yararsızdır. ReLU aktivasyon fonksiyonunun girdisi 2 ise çıktısı da 2 olur. Yani girdileri 0 1 aralığına dönüştürmez. Bir modelin farklı katmanlarında farklı tip aktivasyon fonksiyonları kullanılabilir.

Q3- Assume that you have created an ANN model to find out there is a crack on a concrete or not. The model gives %99 accuracy on the training set and %85 accuracy on the test set. Based on the given information which statement below is correct?

- We need to change activation function to avoid overfitting on training set
- We need to change activation function to avoid underfitting on test set.
- We need to apply regularization to avoid overfitting on test set
- We need to apply regularization to avoid overfitting on training set ✓

A3- Kurulan model training sette overfitting olmuştur. Overfitting durumunu ortadan kaldırmak için modelin eğitimi esnasında regularization uygulamak gerekir.

Q4- Which of the following statements are true about activation functions?

- I- Different layers cannot have different activation functions in a model.
- II- When you use activation functions in your model, you are adding non-linearity.
- III- We are using SoftMax activation function on output layer when we have multi-class model.
- IV- Sigmoid function generates a value between -1 and 1.

- I, II
- II, III ✓
- III, IV
- I, IV
- I, III

A4- Cevap: B

- I- Model içerisindeki farklı layerlar farklı aktivasyon fonksiyonları kullanabilirler. Örneğin iki hidden layer ve 1 output layer bulunduran modelde, ilk hidden layer ReLU, ikinci hidden layer Sigmoid, output layer ise Softmax kullanılabilir.
- II- Aktivasyon fonksiyonları modelde lineer öğrenmeyi kaldırmak için kullanılır.
- III- Softmax fonksiyonu her bir sınıf için olasılıksal bir değer verir. Örneğin resimdeki canlı %80 köpek, %15 kedi, %5 tavşan diyebilir.
- IV- Sigmoid fonksiyonu 0 ile 1 arasında olasılıksal bir değer üretir.

Q5- Which of the following is not one of the evaluation metrics?

- Accuracy
- Recall
- Precision
- F1 score
- Softmax ✓

A5- Accuracy, Recall, Precision ve F1 score modelimizin başarısını değerlendirmek için kullandığımız ölçütlerdendir.

Softmax ise bir aktivasyon fonksiyonudur.

$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$,

$\text{Precision} = \frac{TP}{TP+FP}$,

$\text{Recall} = \frac{TP}{TP+FN}$,

$\text{F1 Score} = 2 * \left(\frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \right)$ formülleri kullanılarak hesaplanırlar.

Q6- Which of the following statements is wrong?

- AUC measures the area underneath the entire ROC curve.
- To model a nonlinear problem, we can directly introduce a nonlinearity with piping each hidden layer node through a nonlinear function.
- Sigmoid activation function converts the weighted sum to a value between 0 and 1.
- AUC ranges in value from -1 to 1. ✓
- Softmax assigns decimal probabilities to each class in a multi-class problem.

A6- AUC değeri 0 ile 1 arasında değişir.

Q7- Which of the following statements are true?

- I- Precision and Recall are basically same things when we are talking about classification.
- II- Precision is equal to the TPR.
- III- Recall is equal to FPR.
- IV- Accuracy is a good metric whether dataset is class imbalanced or not.
- V- Sensitivity and specificity terms are not related with ROC or AUC.

- All of them
- II, III, V
- III, IV
- III, V
- None of them ✓

A7- Hiçbir seçenek doğru değildir.

Q8- We know that there are many ways for backpropagation to become problematic. Which of the following are some examples of it?

- I- Vanishing Gradients.
- II- Exploding Gradients.
- III- Dead Re-LU Units.

- All of them ✓
- I, III
- II, III
- III
- None of them

A8- Tüm maddeler backpropagation'ın problemli olduğu durumlara örnektir.

Q9- Why do we need to use backpropagation in neural networks?

- Forward propagation is slow and it requires computational power.
- Forward propagation is linear so it is insufficient.
- Backpropagation is aimed to decrease cost function. ✓
- Backpropagation uses linear regression techniques so it can solve nonlinear problems.

A9- Yapay sinir ağlarını geri beslemeyle eğitmemizin en nihai amacı maliyet fonksiyonunu azaltmaktır. İlk olarak ileri besleme sonucunda bir maliyet fonksiyonumuzun değeri olur, elde edilen bu değeri azaltmak veya daha doğru bir model oluşturmak amacıyla geri besleme kullanılır.

Q10- Regarding to dropout regularization, which one is correct?

- Dropout regularization works by removing a predefined selection of a fixed number of the units in a network layer.
- It prevents overfitting and provides a way of approximately combining exponentially many different neural network architectures efficiently. ✓
- Dropout consists in randomly setting a fraction rate of input units to 1 at each update during training time, which helps prevent overfitting.
- Dropout can be interpreted as a way of regularizing a neural network by adding noise to its input units.

A10-

- I- Dropout regularization, bir ağ katmanındaki sabit sayıda birimin önceden tanımlanmış değil rastgele tanımlanmış bir seçimini kaldırarak çalışır.
- II- Asıl amacı overfitting i azaltmasıdır.
- III- İşlemlerde yer alması için 1'e değil 0'a set edilmesi gereklidir. 1'e set edilirse, matris çarpımı gibi işlemlere aynı şekilde girer.
- IV- Dropout, giriş birimlerine değil saklı birimlere (hidden units) gürültü ekleyerek bir sinir ağını düzenli hale getirmenin bir yoludur.