Machine Learning Crash Course - Final Test

Toplam puan 100/100

E-posta adresi *

sevdanurgenc@gmail.com

Name-Last Name: *

Sevdanur GENÇ

Q1- Which of the below is the reason why we are using validation sets? * 4/4

- ○ Validation sets help us check if the model overfits or underfits.
- ◉ The regressor might overfit to test set if we don't use validation sets.
- ○ Validation sets increase model's performance on the fit.

Q2- Why and when do we use sparse representation? * 4/4

- ○ When data size is small and a few of feature value that we are interested in is zero.
- ○ When data size is small and a few of feature value that we are interested in is non-zero.
- ○ When data size is large and most of feature value that we are interested in is non-zero.
- ◉ When data size is large and most of feature value that we are interested in is zero.

Q3- Regarding to dropout regularization, which one is correct? * 4/4

- ○ 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
- ◉ Dropout prevents overfitting and provides a way of approximately combining exponentially many different neural network architectures efficiently.
- ○ Dropout regularization works by removing a predefined selection of a fixed number of the units in a network layer

Q4- Which of the following is not a problem caused by not being able to choose the optimal learning rate? * 4/4

- ○ Taking a long step without noticing the minimum point of the loss curve.
- ◉ Taking a step towards the wrong direction for to reach minimum point of the loss curve.
- ○ It takes long time to reach minimum point of the loss curve.

Q5- Which one is true related to the "correlation matrix"? * 4/4

- ○ It indicates how each attribute's feature values relate to the other attributes' raw values.
- ○ The lower the absolute value of a correlation value, the greater its predictive power.
- ○ "0.0" means no correlation; the two columns are linearly correlated.
- ◉ It indicates how each attribute's raw values relate to the other attributes' raw values.

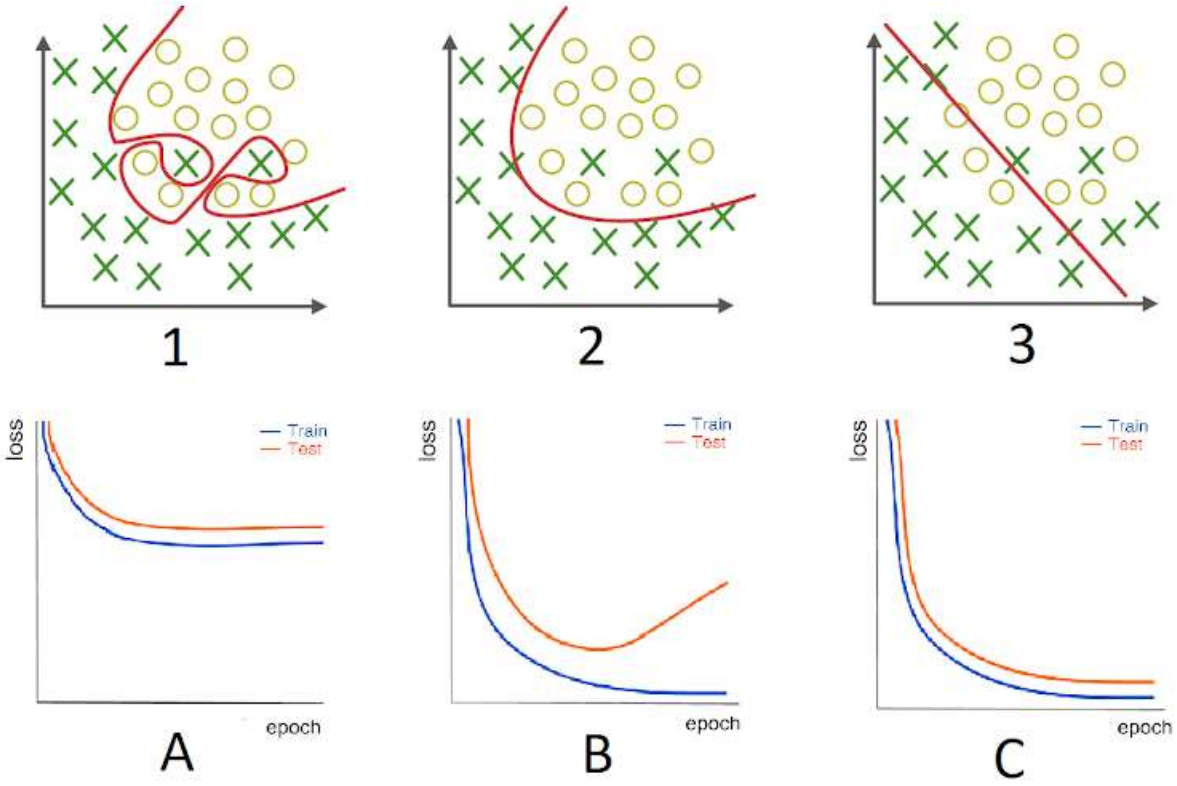
Q6- Which of the following options allows us to create a layer called hidden_layer_2 with 10 neurons and ReLU activation function? * 4/4

- ○ tf.keras.layers.Dense(units='hidden_layer_2', activation='relu',name='10')
- ◉ tf.keras.layers.Dense(units=10, activation='relu',name='hidden_layer_2')
- ○ tf.keras.layers.Dense(units=100, activation='linear',name='hidden_layer_2')
- ○ tf.keras.layers.Dense(units=10, activation='sigmoid',name='hidden_layer_2')

Q7- A dataset with 250 samples is split into training and test set in an 80-20 ratio. The model was trained using Stochastic Gradient Descent using the mini-batch size of 32. After successful 1000 epochs, how many times the model was updated during training? * 4/4

- ○ 200
- ○ 6000
- ◉ 7000
- ○ 32

Q8- You can see different models to classify "X" or "O" for a dataset below. 4/4
Which loss-epoch graph would you expect for these models? *



- [ ] 1-A / 2-B / 3-C
- [ ] 1-B / 2-A / 3-C
- [ ] 1-C / 2-A / 3-B
- [x] 1-B / 2-C / 3-A

Q9- Which of the following statement is false about embeddings? * 4/4

- ○ An Embedding is a translation of a high-dimensional vector into a low-dimensional space.
- ○ An Embedding is a mapping of a categorical variable to a vector of continuous numbers.
- ◉ An Embedding is a sparse vector which has values 0 and 1.
- ○ The main purpose of an embedding is to find nearest neighbors in the embedding space.

Q10- Which of the following statements are true about one-hot encoding? * 4/4

I- One-hot encoding can be used to numeric data that you do not want to directly multiply by a weight, such as serial number of a product.
II- The length of the one-hot encode vector is equal to number of observations in the data.
III- One-hot encoding represents the feature as integers like 0,1,2,3 etc.
IV- Since models cannot learn from categorical data, we are using one-hot encoding in machine learning problems.

- ○ I - III
- ○ III - IV
- ◉ I - IV
- ○ II - IV

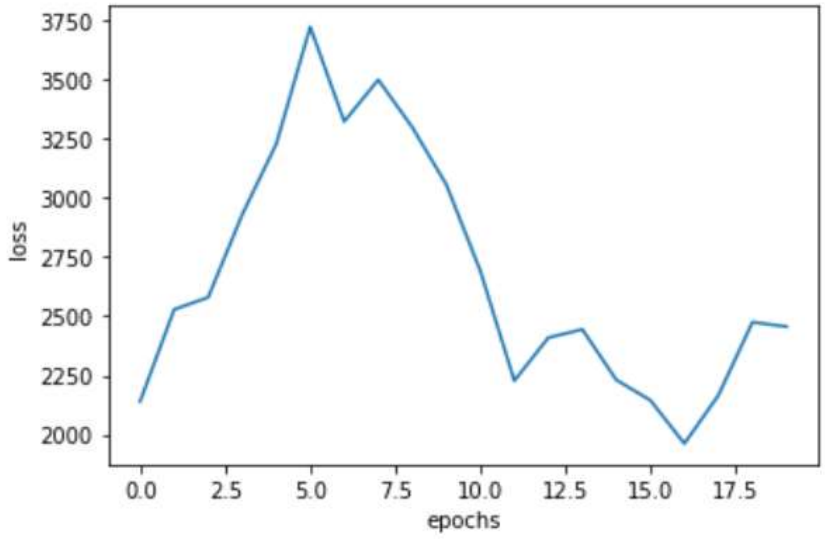
Q11- You have a randomized data that has 10000 instances with 10 feature and 1 class label as X and Y. You want to build a neural network to fit that data with the below code. After fitting you noticed that your loss value is not constantly decreasing like the below graph. What is the reason behind this? * 4/4

The mentioned graph is right after this section.

```
import tensorflow as tf
import numpy as np

X = np.random.rand(10000, 10)
Y = np.random.randint(0,2, size=(10000,))

model = tf.keras.models.Sequential()
model.add(tf.keras.layers.Dense(units=20, input_shape=(10,),
                                activation='relu',
                                name='Hidden1'))
model.add(tf.keras.layers.Dense(units=12,
                                activation='relu',
                                name='Hidden2'))
model.add(tf.keras.layers.Dense(units=1,
                                name='Output'))
model.compile(optimizer=tf.keras.optimizers.SGD(learning_rate=1),
              loss="mean_squared_error",
              metrics=[tf.keras.metrics.MeanSquaredError()])

model.fit(X,Y,epochs=20)
```

- ○ Epoch is too low.
- ○ Epoch is too high.
- ◉ Learning rate is too high.
- ○ Learning rate is too low.
- ○ You can't fit random values to two different classes.

The graph of Q-11



Q12- Which of the following statements is false? * 4/4

- ○ Word2Vec exploits contextual information.
- ○ Sampling bias occurs if proper randomization is not used during data collection.
- ◉ Categorical data refers to input features that represent infinite set of choices.
- ○ Selection bias occurs if a data set's examples are chosen in a way that is not reflective of their real-world distribution.

Q13- Why do we need to use backpropagation in neural networks? * 4/4

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

Q14- Which of the following statement about the feature cross is false? * 4/4

- ○ A feature cross is a synthetic feature that encodes non-linearity in the feature space by multiplying two or more input features together.
- ○ A feature cross might be formed by multiplying the values of two or more features.
- ◉ A feature cross cannot be formed by squaring a single feature.
- ○ Linear learners scale well to massive data. Using feature crosses on massive data sets is one efficient strategy for learning highly complex models.

Q15- Which of the following statements are true for dynamic training? * 4/4

I- Very little or no monitoring of input data need to be done at inference time
II- Very little or no monitoring of training need to be done
III- Our model stays updated as new data arrives
IV- Being stale is not a problem of dynamic training

- ○ I, III
- ○ I, IV
- ◉ III, IV
- ○ II,III

Q16- 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? * 4/4

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

Q17- Below there is an MNIST classifier that trains to 99% accuracy or above, and does it without a fixed number of epochs -- i.e. you should stop training once you reach that level of accuracy. Some notes: - It should succeed in less than 10 epochs, so it is okay to change epochs= to 10, but nothing larger - When it reaches 99% or greater it should print out the string "Reached 99% accuracy so canceling training!" - If you add any additional variables, make sure you use the same names as the ones used in the class. According to this information, which of the below statements numbered green # is false? * 4/4

```
def train_mnist():
    class OurCallback(tf.keras.callbacks.Callback):
        def on_epoch_end(self, epoch, logs={}):
            class OurCallback(tf.keras.callbacks.Callback):
               def on_epoch_end(self, epoch, logs={}):
                  if(logs.get('accuracy')>0.99):   #1
                  print("\nReached 99% accuracy so cancelling training!")  #2
                  self.model.stop_training = True

    mnist = tf.keras.datasets.mnist

    (x_train, y_train),(x_test, y_test) = mnist.load_data(path="mnist.npz")
    x_train = x_train / 255.0
    model = tf.keras.models.Sequential([
        tf.keras.layers.Flatten(),
        tf.keras.layers.Dense(512, activation=tf.nn.relu),
        tf.keras.layers.Dense(10, activation=tf.nn.softmax)    #3
    ])

    model.compile(optimizer='adam',
                  loss='sparse_categorical_crossentropy',
                  metrics=['accuracy'])

    history = model.fit(x_train, y_train, epochs=12, callbacks=[OurCallback()])  #4
    # model fitting
    return history.epoch, history.history['acc'][-1]
```

- ○ #1
- ○ #2
- ○ #3
- ◉ #4

Q18- Which of the following statements are true for static training? * 4/4

I. You can verify the model before applying it in production
II. The model does not stay up to date as new data arrives
III. Static training requires less monitoring of training than dynamic training
IV. We can use batch training and testing, iterate it until model is good

- ○ I, III
- ○ III,IV
- ○ I,III,IV
- ◉ All of them.

Q19- Which of the following statements are true about activation functions? * 4/4

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
- ○ III, IV
- ◉ II,III
- ○ I, IV

Q20- Which of the above options are true about logistic regression? * 4/4

I- Logistic regression is a linear model
II- Logistic regression is a non-linear model
III- Logistic regression generates a probability—a value between 0 and 1
IV- Mean Squared Error is the loss function for logistic regression

- ○ I-IV
- ○ III-IV
- ◉ I-III
- ○ II-III

Q21- Which one should be added to the last layer of the network so the output will be a probability distribution consisting of 25 probabilities. * 4/4

```
model = tf.keras.Sequential([
    feature_layer,
    layers.Dense(150,activation='relu'),
    layers.Dense(100,activation='relu'),
    layers.Dense(50,activation='relu'),
    .....................
])
```

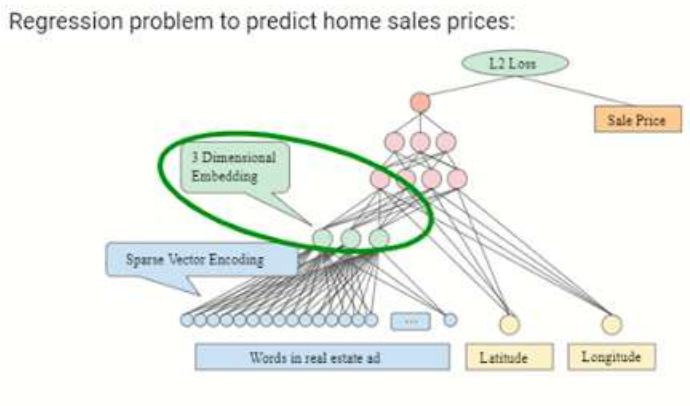
- ○ layers.Dense(100,activation='sigmoid')
- ○ layers.Dense(25,activation='sigmoid')
- ○ layers.Dense(25,activation='relu')
- ◉ layers.Dense(25,activation='softmax')

Q22- 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. * 4/4

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.

- ○ II,III,IV,V,VII
- ◉ II,III,IV,V,VI
- ○ I,II,III,V,VI
- ○ I,II,III,V,VII

Q23- In embeddings, suppose you have 3 units in the embedding layer as shown below. In this case, how would you represent those embeddings and which type of propagation would you use to train the proposed model? * 4/4



- [ ] 3 dimensional representation and back propagation
- [ ] 2 dimensional representation and back propagation
- [ ] 2 dimensional representation and feed forward,back propagation
- [x] 3 dimensional representation and feed forward ,back propagation

Q24- Which of the following statements is wrong? * 4/4

- [ ] AUC measures the area underneath the entire ROC curve.
- [x] AUC ranges in value from -1 to 1.
- [ ] Softmax assigns decimal probabilities to each class in a multi-class problem.
- [ ] To model a nonlinear problem, we can directly introduce a nonlinearity with piping each hidden layer node through a nonlinear function.

Q25- In linear regression with real dataset example, suppose the label is the population as it is defined median_house_value. Which features are the best for predicting the label 'population'? * 4/4

```
# Generate a correlation matrix.
training_df.corr()
```

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population	households	median_income	median_house_value
longitude	1.0	-0.9	-0.1	0.0	0.1	0.1	0.1	-0.0	-0.0
latitude	-0.9	1.0	0.0	-0.0	-0.1	-0.1	-0.1	-0.1	-0.1
housing_median_age	-0.1	0.0	1.0	-0.4	-0.3	-0.3	-0.3	-0.1	0.1
total_rooms	0.0	-0.0	-0.4	1.0	0.9	0.9	0.9	0.2	0.1
total_bedrooms	0.1	-0.1	-0.3	0.9	1.0	0.9	1.0	-0.0	0.0
population	0.1	-0.1	-0.3	0.9	0.9	1.0	0.9	-0.0	-0.0
households	0.1	-0.1	-0.3	0.9	1.0	0.9	1.0	0.0	0.1
median_income	-0.0	-0.1	-0.1	0.2	-0.0	-0.0	0.0	1.0	0.7
median_house_value	-0.0	-0.1	0.1	0.1	0.0	-0.0	0.1	0.7	1.0

- (●) total_rooms,total_bedrooms and households
- ( ) latitude,longitude and housing_median_age
- ( ) median_income,median_house_value and total_rooms
- ( ) households,median_income and total_rooms

Bu içerik Google tarafından oluşturulmamış veya onaylanmamıştır. - Hizmet Şartları - Gizlilik Politikası

Google Formlar