## ▼ CODING: Impact of Training Data Amount

```
# Enable GPU: "Runtime"-->"Change Runtime"-->"Hardware Accelerator"
#1. Check if GPU is enabled
import tensorflow as tf
tf.test.gpu device name()
    '/device:GPU:0'
#2. Import dataset
from keras.datasets import cifar10
import numpy as np
from sklearn.model_selection import train_test_split
(X_train, y_train), (X_test, y_test) = cifar10.load_data()
from tensorflow.keras.utils import to categorical
y_train = to_categorical(y_train)
y_test = to_categorical(y_test)
#X_train, X_test, y_train, y_test = train_test_split(x, y, train_size=0.8)
X train, X val, y train, y val = train test split(X train, y train, test size=0.2)
# ===== MODEL 1 ======== 100% Training Data ===== No Regularization ======
from tensorflow.python.keras import Sequential
from tensorflow.python.keras.layers import Dense, Conv2D, Flatten, MaxPooling2D
model 1 = Sequential()
model_1.add(Conv2D(32, (3, 3), activation='relu', padding ='same', input_shape=(32, 32
model 1.add(MaxPooling2D((2, 2)))
model 1.add(Conv2D(32, (3, 3), activation='relu', padding ='same'))
model 1.add(MaxPooling2D((2, 2)))
model 1.add(Conv2D(64, (3, 3), padding ='same', activation='relu'))
model 1.add(MaxPooling2D((2, 2)))
model 1.add(Conv2D(128, (3, 3), padding ='same', activation='relu'))
model 1.add(MaxPooling2D((2, 2)))
model 1.add(Flatten())
model 1.add(Dense(512, activation='relu'))
model 1.add(Dense(10, activation = "softmax"))
model 1.summary()
model 1.compile(optimizer = 'adam', loss = 'categorical crossentropy', metrics = ['acc
```

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 32, 32, 32)	896
max_pooling2d (MaxPooling2D)	(None, 16, 16, 32)	0
conv2d_1 (Conv2D)	(None, 16, 16, 32)	9248
max_pooling2d_1 (MaxPooling2	(None, 8, 8, 32)	0
conv2d_2 (Conv2D)	(None, 8, 8, 64)	18496
max_pooling2d_2 (MaxPooling2	(None, 4, 4, 64)	0
conv2d_3 (Conv2D)	(None, 4, 4, 128)	73856
<pre>max_pooling2d_3 (MaxPooling2</pre>	(None, 2, 2, 128)	0
flatten (Flatten)	(None, 512)	0
dense (Dense)	(None, 512)	262656
dense_1 (Dense)	(None, 10)	5130
Total params: 370,282 Trainable params: 370,282 Non-trainable params: 0		
# Time how fast the model train import time		
<pre>start = time.time() history = model_1.fit(X_train, y_ end = time.time()</pre>	train, batch_size = 64, e	pochs = 20, verbose = 1, val
<pre>num_mins = (end-start)/60 print("Total training time: " +</pre>	str(num_mins) + " minutes	•")
# Loss and Accuracy		

plt.plot(history.history["loss"], color = "blue", label = "train")

loss, acc = model\_1.evaluate(X\_test, y\_test, verbose =0)

print("Test loss: %.4f" % loss)

from matplotlib import pyplot as plt

# Plot loss function

print("Test accuracy: %.2f" % (acc \* 100.0))

```
plt.plot(history.history["val_loss"], color = "red", label = "test")
plt.legend()
plt.ylabel("Cross entropy loss")
plt.xlabel("Num of epochs")
plt.show()

#1 Plot accuracy

plt.plot(history.history["accuracy"], color = "blue", label = "train")
plt.plot(history.history["val_accuracy"], color = "red", label = "test")
plt.legend()

plt.ylabel("Accuracy")
plt.xlabel("Num of epochs")
plt.show()
```

```
Epoch 1/20
  Epoch 2/20
  Epoch 3/20
  Epoch 4/20
  Epoch 5/20
  Epoch 6/20
  Epoch 7/20
  Epoch 8/20
  Epoch 9/20
  Epoch 10/20
  Epoch 11/20
  Epoch 12/20
  Epoch 13/20
  Epoch 14/20
  Epoch 15/20
  Epoch 16/20
  625/625 [================ ] - 10s 17ms/step - loss: 0.3541 - a
  Epoch 17/20
  Epoch 18/20
  COF / COF -
                      11 ... 10 ... / ... .
# ====== MODEL 2 ========= 100% Training Data ====== With Regularization ======
from tensorflow.python.keras import Sequential
from tensorflow.python.keras.layers import Dense, Conv2D, Flatten, MaxPooling2D
from tensorflow.python.keras.layers.core import Dropout
model 2 = Sequential()
model 2.add(Conv2D(32, (3, 3), activation='relu', padding ='same', input shape=(32, 32
model 2.add(MaxPooling2D((2, 2)))
model 2.add(Dropout(0.1))
model_2.add(Conv2D(32, (3, 3), activation='relu', padding ='same'))
model 2.add(MaxPooling2D((2, 2)))
model 2.add(Dropout(0.1))
model 2.add(Conv2D(64, (3, 3), padding ='same', activation='relu'))
model 2.add(MaxPooling2D((2, 2)))
model 2.add(Dropout(0.1))
model 2.add(Conv2D(128, (3, 3), padding ='same', activation='relu'))
```

```
model_2.add(MaxPooling2D((2, 2)))
model_2.add(Dropout(0.1))

model_2.add(Flatten())
model_2.add(Dense(512, activation='relu'))
model_2.add(Dropout(0.1))
model_2.add(Dense(10, activation = "softmax"))

model_2.summary()
```

model\_2.compile(optimizer = 'adam', loss = 'categorical\_crossentropy', metrics = ['acc

Model: "sequential\_1"

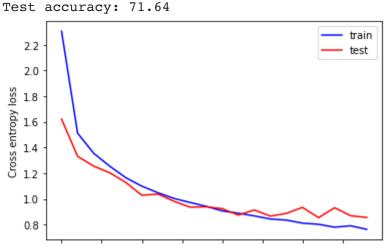
Layer (type)	Output Sh	ape	Param #
conv2d_4 (Conv2D)	(None, 32	32, 32)	896
max_pooling2d_4 (MaxPooling2	(None, 16	, 16, 32)	0
dropout (Dropout)	(None, 16	, 16, 32)	0
conv2d_5 (Conv2D)	(None, 16	, 16, 32)	9248
max_pooling2d_5 (MaxPooling2	(None, 8,	8, 32)	0
dropout_1 (Dropout)	(None, 8,	8, 32)	0
conv2d_6 (Conv2D)	(None, 8,	8, 64)	18496
max_pooling2d_6 (MaxPooling2	(None, 4,	4, 64)	0
dropout_2 (Dropout)	(None, 4,	4, 64)	0
conv2d_7 (Conv2D)	(None, 4,	4, 128)	73856
max_pooling2d_7 (MaxPooling2	(None, 2,	2, 128)	0
dropout_3 (Dropout)	(None, 2,	2, 128)	0
flatten_1 (Flatten)	(None, 51	2)	0
dense_2 (Dense)	(None, 51	2)	262656
dropout_4 (Dropout)	(None, 51	.2)	0
dense_3 (Dense)	(None, 10	)	5130
======================================		:=========	

Total params: 370,282 Trainable params: 370,282 Non-trainable params: 0

<sup>#</sup> Time how fast the model train

```
start = time.time()
history = model_2.fit(X_train, y_train, batch_size = 64, epochs = 20, verbose = 1, val
end = time.time()
num mins = (end-start)/60
print("Total training time: " + str(num_mins) + " minutes.")
# Loss and Accuracy
loss, acc = model_2.evaluate(X_test, y_test, verbose =0)
print("Test loss: %.4f" % loss)
print("Test accuracy: %.2f" % (acc * 100.0))
# Plot loss function
from matplotlib import pyplot as plt
plt.plot(history.history["loss"], color = "blue", label = "train")
plt.plot(history.history["val_loss"], color = "red", label = "test")
plt.legend()
plt.ylabel("Cross entropy loss")
plt.xlabel("Num of epochs")
plt.show()
#1 Plot accuracy
plt.plot(history.history["accuracy"], color = "blue", label = "train")
plt.plot(history.history["val accuracy"], color = "red", label = "test")
plt.legend()
plt.ylabel("Accuracy")
plt.xlabel("Num of epochs")
plt.show()
```

```
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
625/625 [============== ] - 11s 18ms/step - loss: 1.0037 - a
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
625/625 [============= ] - 11s 18ms/step - loss: 0.7801 - a
Epoch 19/20
625/625 [============== ] - 11s 17ms/step - loss: 0.7909 - a
Epoch 20/20
Total training time: 4.406443663438162 minutes.
Test loss: 0.8474
```



## 75% Training Data

```
# ====== 75% Training Data ======

from keras.datasets import cifar10
import numpy as np
from sklearn.model_selection import train_test_split

(X_train, y_train), (X_test, y_test) = cifar10.load_data()

from tensorflow.keras.utils import to_categorical

y_train = to_categorical(y_train)

y_test = to_categorical(y_test)

X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.2)

X_train, X_test75, y_train, y_test75 = train_test_split(X_train, y_train, train_size=0.2)
```

## MODEL 3:75% Training Data | No Regularization

```
from tensorflow.python.keras import Sequential
from tensorflow.python.keras.layers import Dense, Conv2D, Flatten, MaxPooling2D
# Time how fast the model train
start = time.time()
history = model 1.fit(X train, y train, batch size = 64, epochs = 20, verbose = 1, val
end = time.time()
num mins = (end-start)/60
print("Total training time: " + str(num_mins) + " minutes.")
# Loss and Accuracy
loss, acc = model_1.evaluate(X_test, y_test, verbose =0)
print("Test loss: %.4f" % loss)
print("Test accuracy: %.2f" % (acc * 100.0))
# Plot loss function
from matplotlib import pyplot as plt
plt.plot(history.history["loss"], color = "blue", label = "train")
plt.plot(history.history["val_loss"], color = "red", label = "test")
plt.legend()
plt.ylabel("Cross entropy loss")
plt.xlabel("Num of epochs")
```

```
plt.show()
#1 Plot accuracy

plt.plot(history.history["accuracy"], color = "blue", label = "train")
plt.plot(history.history["val_accuracy"], color = "red", label = "test")
plt.legend()

plt.ylabel("Accuracy")
plt.xlabel("Num of epochs")
plt.show()
```

```
Epoch 1/20
 Epoch 2/20
 Epoch 3/20
 Epoch 4/20
 Epoch 5/20
 Epoch 6/20
 Epoch 7/20
 Epoch 8/20
 Epoch 9/20
 Epoch 10/20
 Epoch 11/20
 Epoch 12/20
 Epoch 13/20
 Epoch 14/20
 Epoch 15/20
 Epoch 16/20
MODEL 4: 75% Training Data | With Regularization
 τυν/τυν [------ υσος υ.22υ/ - ας
# Time how fast the model train
start = time.time()
history = model 2.fit(X train, y train, batch size = 64, epochs = 20, verbose = 1, val
end = time.time()
num mins = (end-start)/60
print("Total training time: " + str(num mins) + " minutes.")
# Loss and Accuracy
loss, acc = model 2.evaluate(X test, y test, verbose =0)
print("Test loss: %.4f" % loss)
print("Test accuracy: %.2f" % (acc * 100.0))
# Plot loss function
from matplotlib import pyplot as plt
```

```
plt.plot(history.history["loss"], color = "blue", label = "train")
plt.plot(history.history["val_loss"], color = "red", label = "test")
plt.legend()
plt.ylabel("Cross entropy loss")
plt.xlabel("Num of epochs")
plt.show()

#1 Plot accuracy

plt.plot(history.history["accuracy"], color = "blue", label = "train")
plt.plot(history.history["val_accuracy"], color = "red", label = "test")
plt.legend()

plt.ylabel("Accuracy")
plt.xlabel("Num of epochs")
plt.show()
```

```
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Encah 10/20
ACO/ACO [----- 1 7- 15m-/ston 1-ss. 0 5CO2
```

## 50% Training Data

Test accuracy: 72.65

```
# ====== 50% Training Data ======

from keras.datasets import cifar10
import numpy as np
from sklearn.model_selection import train_test_split

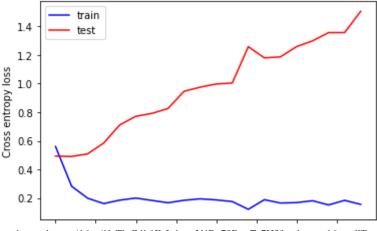
(X_train, y_train), (X_test, y_test) = cifar10.load_data()

from tensorflow.keras.utils import to categorical
```

```
y_train = to_categorical(y_train)
y_test = to_categorical(y_test)
X train, X val, y train, y val = train test split(X train, y train, test size=0.2)
X train, X test75, y train, y test75 = train_test_split(X train, y train, train_size=(
MODEL 5: 50% Training Data | No Regularization
                        /~/ \~/\
# Time how fast the model train
start = time.time()
history = model_1.fit(X_train, y_train, batch_size = 64, epochs = 20, verbose = 1, val
end = time.time()
num mins = (end-start)/60
print("Total training time: " + str(num_mins) + " minutes.")
# Loss and Accuracy
loss, acc = model 1.evaluate(X test, y test, verbose =0)
print("Test loss: %.4f" % loss)
print("Test accuracy: %.2f" % (acc * 100.0))
# Plot loss function
from matplotlib import pyplot as plt
plt.plot(history.history["loss"], color = "blue", label = "train")
plt.plot(history.history["val loss"], color = "red", label = "test")
plt.legend()
plt.ylabel("Cross entropy loss")
plt.xlabel("Num of epochs")
plt.show()
#1 Plot accuracy
plt.plot(history.history["accuracy"], color = "blue", label = "train")
plt.plot(history.history["val accuracy"], color = "red", label = "test")
plt.legend()
plt.ylabel("Accuracy")
plt.xlabel("Num of epochs")
plt.show()
```

```
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
Total training time: 1.7974004904429117 minutes.
Test loss: 3.5814
```

Test loss: 3.5814
Test accuracy: 64.06



```
0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5
Num of epochs
```

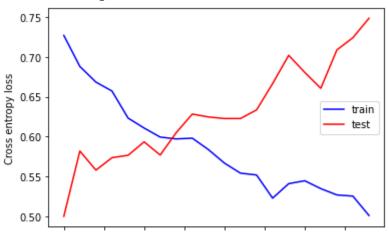
```
0.950 -
0.925 -
0.900 -
0.875 -
0.850 -
```

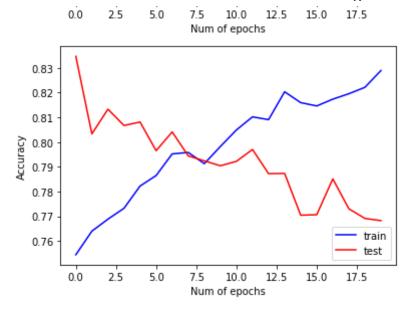
MODEL 6: 50% Training Data | With Regularization

```
0.775 J ____ uain
# Time how fast the model train
start = time.time()
history = model_2.fit(X_train, y_train, batch_size = 64, epochs = 20, verbose = 1, val
end = time.time()
num mins = (end-start)/60
print("Total training time: " + str(num_mins) + " minutes.")
# Loss and Accuracy
loss, acc = model_2.evaluate(X_test, y_test, verbose =0)
print("Test loss: %.4f" % loss)
print("Test accuracy: %.2f" % (acc * 100.0))
# Plot loss function
from matplotlib import pyplot as plt
plt.plot(history.history["loss"], color = "blue", label = "train")
plt.plot(history.history["val loss"], color = "red", label = "test")
plt.legend()
plt.ylabel("Cross entropy loss")
plt.xlabel("Num of epochs")
plt.show()
#1 Plot accuracy
plt.plot(history.history["accuracy"], color = "blue", label = "train")
plt.plot(history.history["val_accuracy"], color = "red", label = "test")
plt.legend()
plt.ylabel("Accuracy")
plt.xlabel("Num of epochs")
plt.show()
```

```
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
Total training time: 2.3696436643600465 minutes.
Test loss: 1.0095
```

Test loss: 1.0095 Test accuracy: 71.75





✓ 3m 24s completed at 11:50 AM