Gurjus Singh

MSDS 458 Artificial Intelligence and Deep Learning

February 7th, 2021

Week 5: A.2 Second Research/Programming Assignment

Abstract

In this research I explored the CIFAR 10 Dataset which was used to train Convolutional Neural Networks and Deep Neural Networks. The CIFAR Dataset consists of 60,000 images of 10 classes. This was used for multiclass classification supervised learning. In the assignment I did a total of 15 experiments and compared the results of each experiment. I also did T-SNDE on the best model and explored how Convolutional Neural Networks learn.

Introduction

In this research that I conducted; I used the CIFAR10 dataset. The dataset consists ten classes labeled airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. With this dataset the goal was to experiment using Convolutional Neural Networks and do multiclass classification. The dataset is composed of 60,000 images, and I wanted to see the outputs of the layers of the Convolutional Neural Network as well as perform T-SNDE. After experimenting CNNs I then compared each Neural Network and made a management recommendation.

Literature Review

Several Researchers have used the CIFAR-10 Dataset. The data was collected by Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton [1].

One study done by Akwasi Darkwah Akwaboah titled “Convolutional Neural Network for CIFAR-10 Dataset Image Classification” explores the use of CNNs in classification of

CIFAR-10 dataset [2]. Akwaboah mentions Alex Krizhevsky who developed the AlexNet architecture which obtained record performance metrics for the CIFAR-10 dataset [2]. He then explains the magic behind CNNs saying that they extract “higher level representation of image features” [2]. They are also good at narrowing network parameters [2] .

The researcher went on to play around with CNNs with the CIFAR 10 dataset. He made 3 Convolutional networks and 2 of them were prone to overfitting [2]. For this third experiment he used L2 Regularizer and dropout of 50 percent and trained on more epochs to 40 to raise accuracy which increased to 75 percent to test data [2].

Methods

In this research I first started out by importing packages such as Sklearn for providing metrics for each model, Tensorflow and Keras for making the CNNs and DNNs. I also imported Numpy for using arrays, pandas for DataFame purposes, and matplotlib, seaborn for data visualizations. The packages used are seen in 1-1.

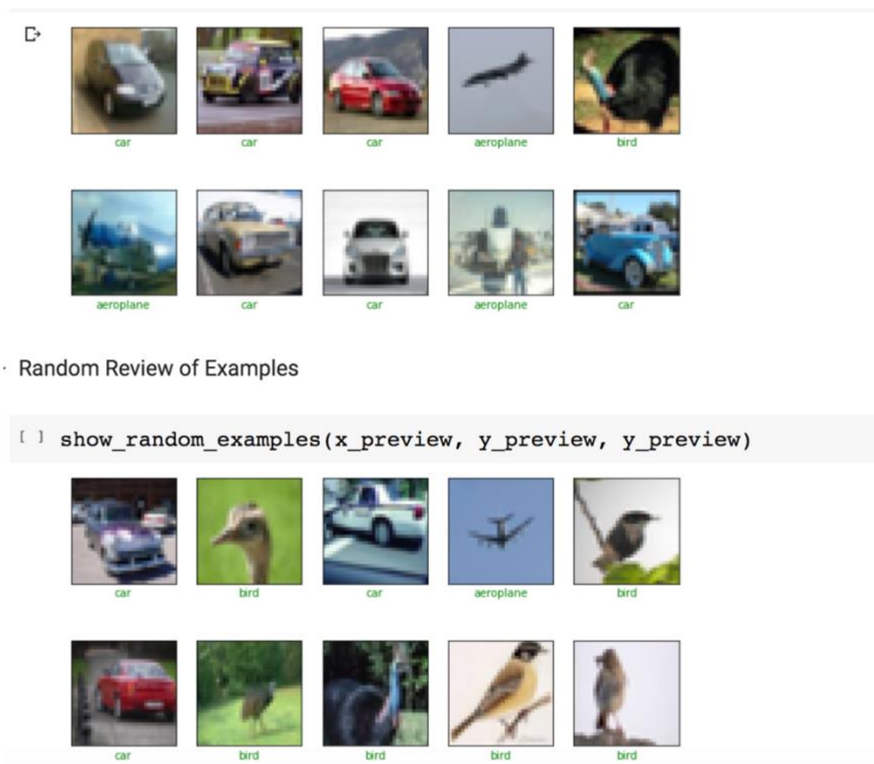
```
# Helper libraries
import datetime
from packaging import version
import matplotlib.pyplot as plt
import seaborn as sns

import sklearn
from sklearn.metrics import confusion_matrix
from collections import Counter
import numpy as np
import pandas as pd

# TensorFlow and tf.keras
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras.utils import to_categorical
from tensorflow.keras import models, layers
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D
from tensorflow.keras.layers import Dropout, Flatten, Input, Dense
```

Import Packages 1-1

I first checked the versions of the Tensorflow packages to make sure I was using 2.0 or above of Tensorflow. Next, I downloaded the CIFAR-10 dataset and split it into train, and test. I check out how many images in Train and Test which were 50,000 and 10,000 respectively. I also looked at the labels to see what they were of. I then plotted some of the pictures and got their labels as seen in 1-2. After I looked at the dataset, I normalized the dataset pixels by dividing by 255 so each pixel value was between 0 and 1. Once I normalized the dataset, I split up the dataset further into the validation set. The validation set contained 10,000 images.

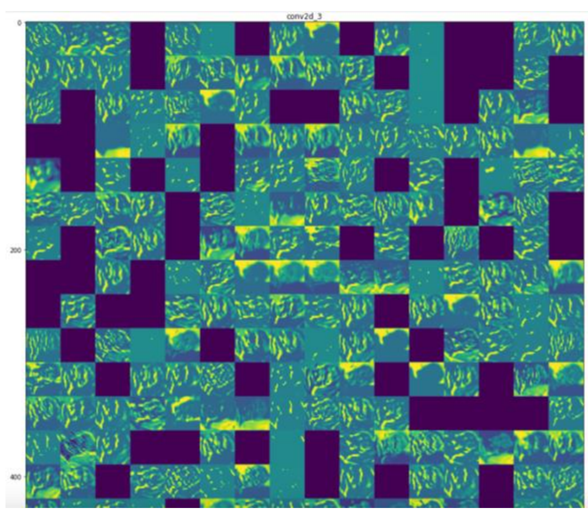


Picture Previews 1-2

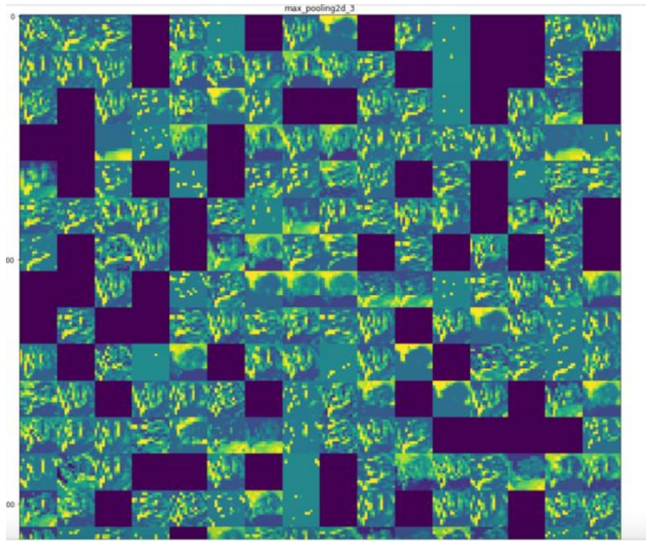
Once I was done it was time to create each model and experiment. I created a total of 15 experiments. I used techniques such as L1 and L2 regularizers, dropout and early stopping which are all considered regularization techniques. What was special about this research was that I created CNNs which is good at finding features from images. CNNs involve max pooling and convolution layers. Max pooling help in reducing the number of parameters in the network by "downsizing feature maps" [3]. The convolutional layer functions in that tries to extract local patterns from a window [3]. The window is called a filter which goes over each part of the image extracting the features in the image [3].

For the models I used 64 neurons for my first two experiments using Deep Neurla Networks while my 3rd and 4th experiments were CNNs where there were 2 Max Pooling/Convolution Layers and 3 Max Pooling/Convolution Layers respectively. I then applied regularization techniques such as Early Stopping, L1/L2 Regularizers, and Dropout after the first 4 experiments.

Later after implementing the models I took a look at the accuracies to find the best model. I believe that Experiment 16 performed the best by looking at the results. I deep dived into Experiment 16 by also looking at other metrics such as Confusion Matrix and Precision Score. I also performed T-SNDE on this model shown in 1-3. For Experiment 3, I wanted to know how the CNN worked so I extracted the outputs shown in 1-3, 1-4. What I got out of 1-3, 1-4 was it was extracting stripe like features which were diagonal lines.



Output of Convolution Layer Experiment 3 1-3



Output of Max Pooling Layer Experiment 3 1-4

Results

EXPERIMENTS - 20 EPOCHS	TRAIN ACCURACY	VAL ACCURACY	TEST ACCURACY	TRAIN LOSS	VAL LOSS	TEST LOSS	TIME (SECS)
1 - 64 NEURONS	60.26%	46.50%	46.98%	1.1515	1.5497	1.5329	52.2
2 - 64 NEURONS	69.13%	46.39%	46.04%	0.9137	1.733	1.7246	66.77
3 - 2 MAX/CONV LAYERS	97%	72.41%	71.30%	1.06E-01	1.1657	1.2416	277.33
4 - 3 MAX/CONV LAYERS	97%	75.24%	73.61%	9.57E-02	1.1107	1.1943	326
5 - EARLY STOPPING - ; SAME AS EXP. 1	55%	49.33%	48.49%	1.3178	1.4717	1.4673	21.51
6 - EARLY STOPPING -	55%	49.08%	49.28%	1.3002	1.4606	1.4447	24.63
7 - EARLY STOPPING - SAME AS EXP 3	91.11%	72.04%	71.84%	0.2771	0.978	1.0158	187.68
8 - EARLY STOPPING - SAME AS EXP 4	96.31%	74.22%	73.94%	0.1168	1.0574	1.1024	300.588
9 - L1L2 REGULARIZER - - 128 NEURONS 2 LAYERS	10.16%	9.37%	10.00%	3.7578	3.7639	3.7637	81.98
10 - L1L2 REGULARIZER -128 NEURONS 3 LAYERS	10.21%	9.37%	10.00%	3.782	3.7852	3.785	117.72
11 - 2 MAX/CONV LAYERS; L1L2 REGULARIZERS	43.96%	43.83%	43.59%	23.7165	23.4212	23.4321	279.39
12 - 3 MAX/CONV LAYERS; ; L1L2 REGULARIZERS	55.85%	52.69%	53.31%	3.7741	3.7843	3.7923	325.26
13 - 512 NEURONS 2 LAYERS; 20 % DROPOUT TECHNIQUE	64.85%	46.81%	47.24%	1.0396	1.6045	1.5765	96.54
14 - 512 NEURONS 3 LAYERS; 20% DROPOUT TECHNIQUE	76.04%	44.16%	44.64%	0.7119	0.7604	1.9415	130.76
15 - 2 MAX/CONV LAYERS; 20% DROPOUT	96.42%	74.24%	73.58%	0.1147	0.9646	1.031	279.16
16 - 3 MAX/CONV LAYERS; 20% DROPOUT	93.70%	75.97%	75.35%	0.1782	0.9382	0.9702	327.04

Results from 16 Experiments 1-5

I ran 16 experiments, where in the first 4 experiments I did not use regularization. I also ran 20 epochs for each experiment. Experiment One involved 64 neurons, Two layers, Experiment Two involved 64 neurons Three layers, Experiment Three involved Two Max Pooling and Convolution Layers, while Experiment Four involved Three Max Pooling and Convolution Layers. After the first four experiments, I then increased neurons involved and use regularization such as L1-L2 Regularizers, Dropout, and Early Stopping. One can see the results and architecture from 1-5. After running each model, I compared the results, and found

Experiment 16 with 20 percent Dropout and 3 Max Pooling, Convolution Layers to be the best model. It had a 93.7 percent Training Accuracy, and 75.35 percent Testing Accuracy. One can see from this that the model was still overfitting, but I examined the other 15 experiments and saw all of them were either overfitting, or underfitting, so I chose a model that was overfitting the least which was Experiment 16.

After choosing what I thought was the best model, I then also looked at the Train, Test Confusion Matrix. The Confusion Matrix in 1-6 shows that the Training Data has the higher accuracy, compared to 1-7 which balanced columns which means that accuracy is less. When the Confusion Matrix has more numbers on the diagonals it means that the model is performing well.

```
[177] conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
      conf_mx

<tf.Tensor: shape=(10, 10), dtype=int32, numpy=
array([[3931,    0,   18,    6,    3,    2,    2,    1,   23,    9],
       [   1, 3992,    0,    3,    0,    1,    0,    0,   18,   11],
       [  21,    0, 3811,   31,   43,   42,   11,    3,    4,    2],
       [   3,    0,   11, 3813,   25,   97,   14,    7,    9,    5],
       [   9,    0,   13,   17, 3945,    5,    4,    2,    6,    0],
       [   1,    1,    2,   60,   27, 3959,    1,    9,    2,    1],
       [   2,    0,   13,   13,   18,   14, 3902,    2,    4,    2],
       [   0,    0,    0,    2,    8,    9,    0, 3978,    0,    2],
       [   4,    1,    0,    2,    1,    0,    0,    1, 3963,    3],
       [   3,    0,    1,    6,    1,    1,    0,    2,   12, 3993]],
      dtype=int32)>
```

Training Confusion Matrix 1-6

```
[179] conf_mx = tf.math.confusion_matrix(test_labels, pred_classestest)
      conf_mx

<tf.Tensor: shape=(10, 10), dtype=int32, numpy=
array([[786,   8,  41,  14,  22,   9,   8,   9,  71,  32],
       [ 19, 863,   3,   8,   2,   4,   6,   2,  30,  63],
       [ 54,   5, 611,  76,  85,  76,  44,  20,  18,  11],
       [ 15,   4,  47, 572,  71, 172,  46,  35,  19,  19],
       [ 18,   2,  42,  44, 761,  41,  29,  46,  13,   4],
       [ 14,   5,  23, 149,  48, 696,  10,  45,   3,   7],
       [  8,   8,  35,  59,  35,  25, 811,   8,   8,   3],
       [ 21,   1,  14,  32,  58,  70,   2, 778,   6,  18],
       [ 49,  14,   8,  13,   2,   3,   4,   2, 875,  30],
       [ 19,  64,   8,  10,   7,   7,   3,   8,  29, 845]], dtype=int32)>
```

Testing Confusion Matrix 1-7

After looking at the Confusion Matrix I also looked at the plot which shows the loss and accuracy as seen in 1-8. I saw that there was a huge gap between Train Loss and Val Loss, it shows that the experiment is overfitting. I also looked at the precision score and it showed that the Precision score for Training Data was significantly higher than Testing Data seen in 1-9. Lastly, I performed T-SNE which is dimensionality reduction unsupervised learning technique to visualize higher dimension data as seen in 1-10 [4].

1-8 Experiment 16 Plot

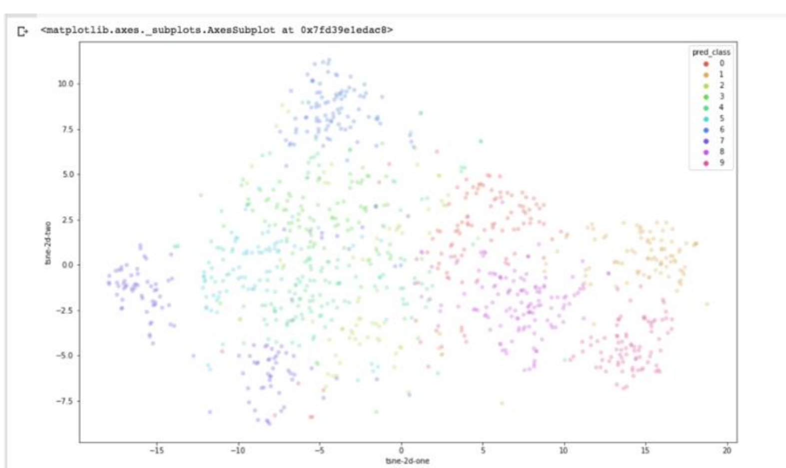
```
[412] precision_score(train_labels, pred_classes,  average='micro')

0.959075

precision_score(test_labels, pred_classestest,  average='micro')

0.7314
```

Precision Score 1-9



1-10 T-SNE

Conclusion

In this research I used CIFAR-10 data to do supervised learning multiclass classification. I explored using CNNs to do the experiments. I then examined the accuracies for each of the experiments and also examined how the Convolution and Max Pooling layers learn. After exploring CNNs. I chose the best experiment which was Experiment 16 and examined the confusion matrices, precision scores, and the loss, accuracy plots. After examining the metrics, I then performed dimensionality reduction with T-SNE. After doing this research **for my management recommendation, I choose the Experiment 16 which was the 3 Max/Conv Layers with 20% Dropout architecture.** I chose this model as it had a higher accuracy than the other experiments with accuracies 93.7 and 75.35 percent as seen in 1-5, and it also overfitted less. Overall, my expectations were lowered for this research as I noticed the accuracies were not that high. I was extremely disappointed in my results, and for further research would like to try to obtain 85-90 percent Testing accuracy, by examining other ways to improve the models.

References

[1] CIFAR-10 and CIFAR-100 datasets. (n.d.). CS Toronto. Retrieved February 5, 2021, from https://www.cs.toronto.edu/%7Ekriz/cifar.html

[2] Akwaboah, A. D. (2019, November). *Convolutional Neural Network for CIFAR-10 Dataset Image Classification*. Research Gate. https://www.researchgate.net/publication/337240963_Convolutional_Neural_Network_for_CIFAR-10_Dataset_Image_Classification

[3] Chollet, F. (2017). *Deep Learning with Python*. Manning Publications Company.

[4] Introduction to t-SNE. (n.d.). Data Camp. Retrieved February 6, 2021, from https://www.datacamp.com/community/tutorials/introduction-t-sne

Appendix

MSDS458 Research Assignment 2

More Technical: Throughout the notebook. This types of boxes provide more technical details and extra references about what you are seeing. They contain helpful tips, but you can safely skip them the first time you run through the code.

The CIFAR-10 dataset (Canadian Institute For Advanced Research) is a collection of images that are commonly used to train machine learning and computer vision algorithms. It is one of the most widely used datasets for machine learning research. The CIFAR-10 dataset contains 60,000 32x32 color images in 10 different classes. The 10 different classes represent airplanes, cars, birds, cats, deer, dogs, frogs, horses, ships, and trucks. There are 6,000 images of each class.

Import packages needed

```
# Helper libraries
import datetime
from packaging import version
import matplotlib.pyplot as plt
import seaborn as sns

import sklearn
from sklearn.metrics import confusion_matrix
from collections import Counter
import numpy as np
import pandas as pd

# TensorFlow and tf.keras
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras.utils import to_categorical
from tensorflow.keras import models, layers
```

```
from tensorflow.keras import models, layers
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D
from tensorflow.keras.layers import Dropout, Flatten, Input, Dense

%matplotlib inline
np.set_printoptions(precision=3, suppress=True)
```

Verify TensorFlow Version and Keras Version

```
print("This notebook requires TensorFlow 2.0 or above")
print("TensorFlow version: ", tf.__version__)
assert version.parse(tf.__version__).release[0] >=2
```

```
This notebook requires TensorFlow 2.0 or above
TensorFlow version:  2.4.1
```

```
print("Keras version: ", keras.__version__)
```

```
Keras version:  2.4.0
```

Suppress warning messages

```
def warn(*args, **kwargs):
    pass
import warnings
warnings.warn = warn
```

Mount Google Drive to Colab Enviorment

```
from google.colab import drive
drive.mount('/content/gdrive')
```

```
Mounted at /content/gdrive
```

Loading cifar10 Dataset

The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly-selected images from each class. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class.

```
(train_images, train_labels),(test_images, test_labels)= tf.keras.
```

```
Downloading data from https://www.cs.toronto.edu/~kriz/cifar-10-python.tar.gz
170500096/170498071 [==============================] - 2s 0us/step
```

- Tuple of Numpy arrays: (x_train, y_train), (x_test, y_test).
- x_train, x_test: uint8 arrays of color image data with shapes (num_samples, 32, 32).
- y_train, y_test: uint8 arrays of digit labels (integers in range 0-9)

EDA Training and Test Datasets

- Imported 50000 examples for training and 10000 examples for test
- Imported 50000 labels for training and 10000 labels for test

```
print('train_images:\t{}'.format(train_images.shape))
print('train_labels:\t{}'.format(train_labels.shape))
print('test_images:\t\t{}'.format(test_images.shape))
print('test_labels:\t\t{}'.format(test_labels.shape))
```

```
train_images:	(50000, 32, 32, 3)
train_labels:	(50000, 1)
test_images:		(10000, 32, 32, 3)
test_labels:		(10000, 1)
```

Review labels for training dataset

```
print("First ten labels training dataset:\n {}\n".format(train_lab
print("This output the numeric label, need to convert to item desc
```

```
First ten labels training dataset:
 [[6]
 [9]
 [9]
 [4]
 [1]
 [1]
 [2]
 [7]
 [8]
 [3]]

This output the numeric label, need to convert to item description
```

▾ Plot Examples

```
def get_three_classes(x, y):
    def indices_of(class_id):
        indices, _ = np.where(y == float(class_id))
        return indices

    indices = np.concatenate([indices_of(0), indices_of(1), indice

    x = x[indices]
    y = y[indices]

    count = x.shape[0]
    indices = np.random.choice(range(count), count, replace=False)

    x = x[indices]
    y = y[indices]

    y = tf.keras.utils.to_categorical(y)

    return x, y

(x_train, y_train), (x_test, y_test) = tf.keras.datasets.cifar10.l

x_preview, y_preview = get_three_classes(x_train, y_train)
x_preview, y_preview = get_three_classes(x_test, y_test)

class_names_preview = ['aeroplane', 'car', 'bird']

def show_random_examples(x, y, p):
    indices = np.random.choice(range(x.shape[0]), 10, replace=Fals

    x = x[indices]
    y = y[indices]
    p = p[indices]

    plt.figure(figsize=(10, 5))
    for i in range(10):
        plt.subplot(2, 5, i + 1)
        plt.imshow(x[i])
```

```
        plt.xticks([])
        plt.yticks([])
        col = 'green' if np.argmax(y[i]) == np.argmax(p[i]) else '
        plt.xlabel(class_names_preview[np.argmax(p[i])], color=col
    plt.show()
```

```
show_random_examples(x_preview, y_preview, y_preview)
```

Random Review of Examples

```
show_random_examples(x_preview, y_preview, y_preview)
```

Preprocessing Data for Model Development

The labels are an array of integers, ranging from 0 to 9. These correspond to the class of clothing the image represents:

Label	Class_
0	airplane
1	automobile
2	bird
3	cat
4	deer
5	dog
6	frog
7	horse
8	ship
9	truck

```
class_names = [['airplane'
,'automobile'
,'bird'
,'cat'
,'deer'
,'dog'
,'frog'
,'horse'
,'ship'
,'truck']]
```

Preprocessing the Examples

The images are 28x28 NumPy arrays, with pixel values ranging from 0 to 255.

1. Each element in each example is a pixel value
2. Pixel values range from 0 to 255
3. 0 = black
4. 255 = white

```
train_images_norm = train_images.astype('float32')/255.
test_images_norm = test_images.astype('float32')/255.
```

```
train_images_norm.shape, test_images_norm.shape
```

```
((50000, 32, 32, 3), (10000, 32, 32, 3))
```

Validating our approach

10,000 samples of our training data to use as a validation set.

```
val_images_norm, train_images_norm = train_images_norm[:10000], tr
val_labels, train_labels = train_labels[:10000], train_labels[1000
```

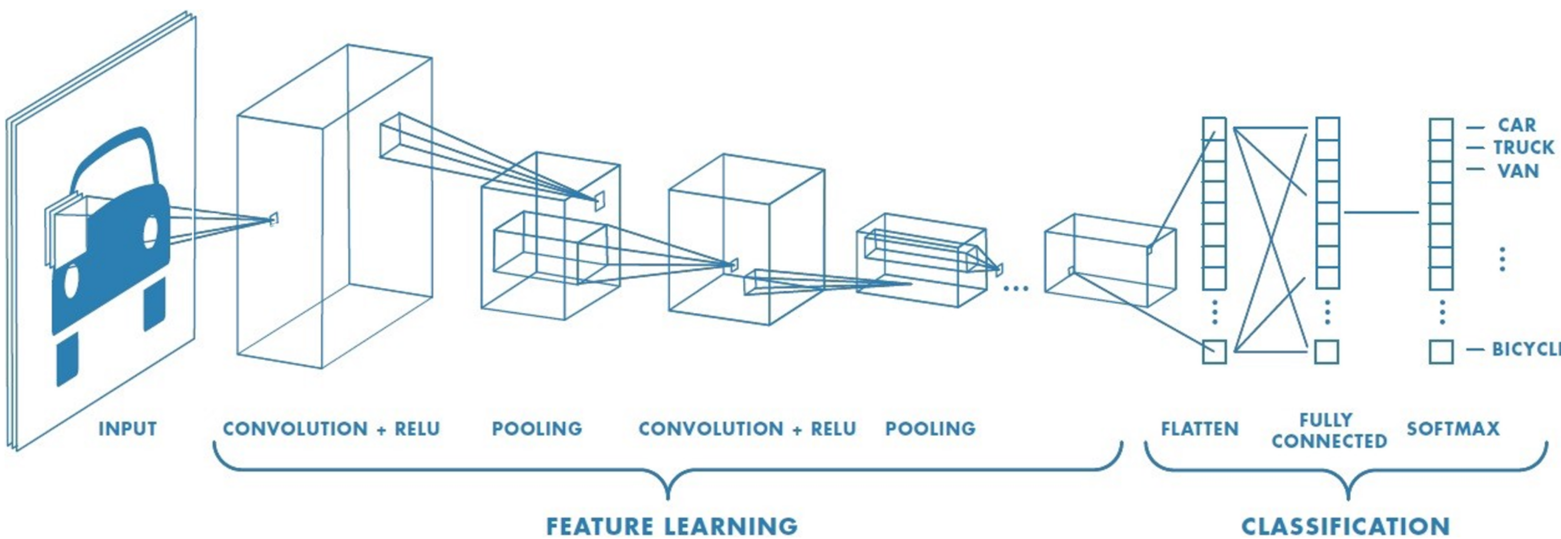
```
val_images_norm.shape, val_labels.shape
```

```
((10000, 32, 32, 3), (10000, 1))
```

```
train_images_norm.shape, train_labels.shape
```

```
((40000, 32, 32, 3), (40000, 1))
```

Create the Model



Build CNN Model

We use a Sequential class defined in Keras to create our model. The first 4 layers Conv2D and MaxPooling handle feature learning. The last 3 layers, handle classification.

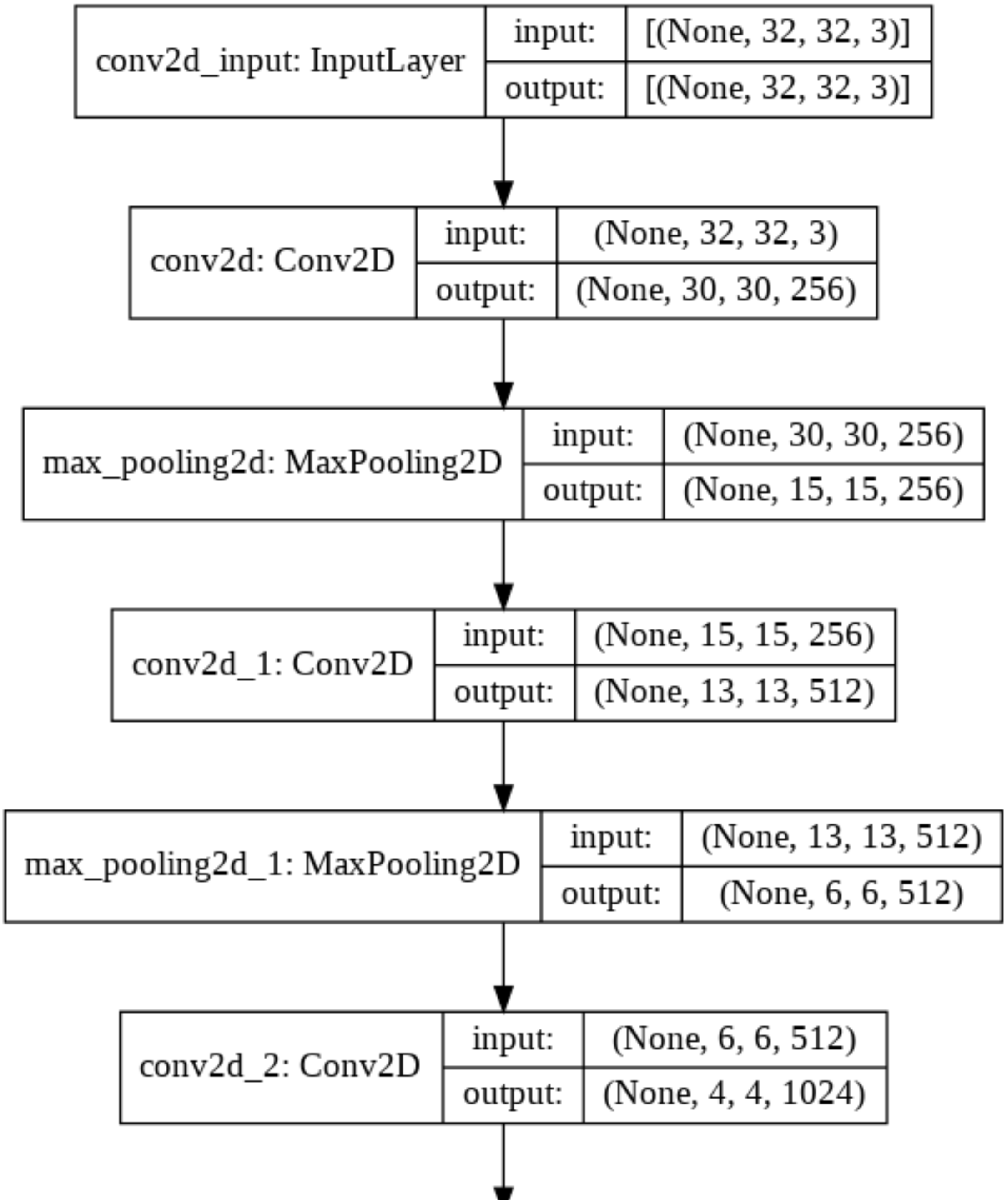
```
model = models.Sequential()
model.add(layers.Conv2D(filters=256, kernel_size=(3, 3), strides=(
model.add(layers.MaxPooling2D((2, 2),strides=2))
model.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides=(
model.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), strides=
model.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model.add(layers.Flatten())
```

```
model.add(layers.Dense(units=512, activation=tf.nn.relu))
model.add(layers.Dense(units=10, activation=tf.nn.softmax))
```

```
model.summary()
```

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
conv2d (Conv2D)              (None, 30, 30, 256)       7168      
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 15, 15, 256)       0         
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 13, 13, 512)       1180160   
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 6, 6, 512)         0         
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 4, 4, 1024)        4719616   
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 2, 2, 1024)        0         
_________________________________________________________________
flatten (Flatten)            (None, 4096)              0         
_________________________________________________________________
dense (Dense)                (None, 512)               2097664   
_________________________________________________________________
dense_1 (Dense)              (None, 10)                5130      
=================================================================
Total params: 8,009,738
Trainable params: 8,009,738
Non-trainable params: 0
_________________________________________________________________
```

```
keras.utils.plot_model(model, "CIFAR10.png", show_shapes=True)
```



Compiling the model

In addition to setting up our model architecture, we also need to define which algorithm should the model use in order to optimize the weights and biases as per the given data. We will use stochastic gradient descent.

We also need to define a loss function. Think of this function as the difference between the predicted outputs and the actual outputs given in the dataset. This loss needs to be minimised in order to have a higher model accuracy. That's what the optimization algorithm essentially does - it minimises the loss during model training. For our multi-class classification problem, categorical cross entropy is commonly used.

Finally, we will use the accuracy during training as a metric to keep track of as the model trains.

tf.keras.losses.SparseCategoricalCrossentropy
https://www.tensorflow.org/api_docs/python/tf/keras/losses/SparseCategoricalCrossentropy

```
model.compile(optimizer='adam',
```

```
          loss=tf.keras.losses.SparseCategoricalCrossentropy(f
          metrics=['accuracy'])
```

Training the model

Module: tf.keras.callbacks

Double-click (or enter) to edit

tf.keras.callbacks.EarlyStopping
https://www.tensorflow.org/api_docs/python/tf/keras/callbacks/EarlyStopping

tf.keras.callbacks.ModelCheckpoint
https://www.tensorflow.org/api_docs/python/tf/keras/callbacks/ModelCheckpoint

```
fit(train_images_norm
    ,train_labels
    ,epochs=20
    ,batch_size=512
    ,validation_data=(val_images_norm,val_labels)
    ,callbacks=[
    tf.keras.callbacks.EarlyStopping(monitor='val_accuracy', patie
    tf.keras.callbacks.ModelCheckpoint('/content/gdrive/My Drive/C
                                save_weights_only=False, monitor='val_accu
   )
```

```
Epoch 1/20
48/79 [=================>............] - ETA: 2s - loss: 2.1804 - accuracy: 0.1916
---------------------------------------------------------------------------
KeyboardInterrupt                         Traceback (most recent call last)
<ipython-input-27-385a9a8be1a5> in <module>()
      7                  tf.keras.callbacks.EarlyStopping(monitor='val_accuracy',
patience=2),
      8                  tf.keras.callbacks.ModelCheckpoint('/content/gdrive/My
Drive/Colab Notebooks/model_{val_accuracy:.4f}.h5', save_best_only=True,
----> 9                                                 save_weights_only=False,
monitor='val_accuracy')]
     10                 )

12 frames
/usr/local/lib/python3.6/dist-packages/tensorflow/python/framework/ops.py in _numpy(self)
   1035   def _numpy(self):
   1036     try:
-> 1037       return self._numpy_internal()
   1038     except core._NotOkStatusException as e:  # pylint: disable=protected-access
   1039       six.raise_from(core._status_to_exception(e.code, e.message), None)  #
pylint: disable=protected-access
```

Validation Data

Data on which to evaluate the loss and any model metrics at the end of each epoch
The model will not be trained on this data

Evaluate the model

In order to ensure that this is not a simple "memorization" by the machine, we should evaluate the performance on the test set. This is easy to do, we simply use the `evaluate` method on our model.

```
loss, accuracy = model.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)
```

Predictions

```
preds = model.predict(test_images_norm)
print('shape of preds: ', preds.shape)
```

Plotting Performance Metrics

We use Matplotlib to create 2 plots--displaying the training and validation loss (resp. accuracy) for each (training) epoch side by side.

```
history_dict = history.history
history_dict.keys()
```

```
history_df=pd.DataFrame(history_dict)
history_df.tail()
```

```
losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)
```

```
plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
```

```
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()
```

Creating confusion matrices

Let us see what the confusion matrix looks like. Using both `sklearn.metrics`. Then we visualize the confusion matrix and see what that tells us.

Get the predicted classes

```
pred_classes = np.argmax(model.predict(train_images_norm), axis=-1
pred_classes
```

Visualizing the confusion matrix

```
conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

```
plt.figure(figsize=(16,8))
plt.matshow(conf_mx, cmap=plt.cm.Blues,fignum=1)
plt.xlabel("Predicted Classes")
plt.ylabel("Actual Classes")
plt.show()
```

Load HDF5 Model Format

tf.keras.models.load_model
https://www.tensorflow.org/api_docs/python/tf/keras/models/load_model

```
model = tf.keras.models.load_model('/content/gdrive/My Drive/Colab
```

```
preds = model.predict(test_images_norm)
```

```
preds.shape
```

```
print("The first predictions\n {}\n".format(preds[0]))
```

```
print(class_names)

print("First ten entries of the predictions:\n {}\n".format(preds[
```

Predictions

```
cm = sns.light_palette((260, 75, 60), input="husl", as_cmap=True)

df = pd.DataFrame(preds[0:20], columns = ['airplane', 'automobile'
df.style.format("{:.2%}").background_gradient(cmap=cm)
```

EXPERIMENT 1

```
model1 = models.Sequential()
model1.add(layers.Dense(units=64, activation=tf.nn.relu))
model1.add(layers.Dense(units=64, activation=tf.nn.relu))
model1.add(layers.Flatten())
model1.add(layers.Dense(units=10, activation=tf.nn.softmax))

model1.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

from time import perf_counter
time = perf_counter()
history1 = model1.fit(train_images_norm
                      ,train_labels
                      ,epochs=20
                      ,batch_size=512
                      ,validation_data=(val_images_norm,val_labels)
                     )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model1.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)
```

```
history_dict = history1.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()
```

```
losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)
```

```
plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()
```

```
pred_classes = np.argmax(model1.predict(train_images_norm), axis=-
```

```
conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

Experiment 2

```
model2 = models.Sequential()
model2.add(layers.Dense(units=64, activation=tf.nn.relu))
model2.add(layers.Dense(units=64, activation=tf.nn.relu))
model2.add(layers.Dense(units=64, activation=tf.nn.relu))
model2.add(layers.Flatten())
model2.add(layers.Dense(units=10, activation=tf.nn.softmax))
```

```
model2.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])
```

```
time = perf_counter()
history2 = model2.fit(train_images_norm
```

```
history2 = model2.fit(train_images_norm
                      ,train_labels
                      ,epochs=20
                      ,batch_size=512
                      ,validation_data=(val_images_norm,val_labels)
                     )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model2.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history2.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()

losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()

pred_classes = np.argmax(model2.predict(train_images_norm), axis=-

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

Experiment 3

```
model3 = models.Sequential()
model3.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides=
model3.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model3.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), strides
model3.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model3.add(layers.Flatten())
model3.add(layers.Dense(units=512, activation=tf.nn.relu))
model3.add(layers.Dense(units=10, activation=tf.nn.softmax))

model3.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

from time import perf_counter
time = perf_counter()
history3 = model3.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model3.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history3.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()

losses = history3.history['loss']
accs = history3.history['accuracy']
val_losses = history3.history['val_loss']
val_accs = history3.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
```

```
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()

pred_classes = np.argmax(model3.predict(train_images_norm), axis=-

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx

from keras.preprocessing import image
import numpy as np



img_tensor = image.img_to_array(test_images[2004])
img_tensor = np.expand_dims(img_tensor, axis=0)
# Remember that the model was trained on inputs
# that were preprocessed in the following way:
img_tensor /= 255.


from keras import models

# Extracts the outputs of the top 8 layers:
layer_outputs = [layer.output for layer in model3.layers[:2]]
# Creates a model that will return these outputs, given the model
activation_model = models.Model(inputs=model3.input, outputs=layer

# This will return a list of 5 Numpy arrays:
# one array per layer activation
activations = activation_model.predict(img_tensor)

first_layer_activation = activations[0]
print(first_layer_activation.shape)


import keras

# These are the names of the layers, so can have them as part of o
```

```
layer_names = []
for layer in model3.layers[:2]:
    layer_names.append(layer.name)

images_per_row = 8

# Now let's display our feature maps
for layer_name, layer_activation in zip(layer_names, activations):
    # This is the number of features in the feature map
    n_features = layer_activation.shape[-1]

    # The feature map has shape (1, size, size, n_features)
    size = layer_activation.shape[1]

    # We will tile the activation channels in this matrix
    n_cols = n_features // images_per_row
    display_grid = np.zeros((size * n_cols, images_per_row * size))

    # We'll tile each filter into this big horizontal grid
    for col in range(n_cols):
        for row in range(images_per_row):
            channel_image = layer_activation[0,
                                             :, :,
                                             col * images_per_row
            # Post-process the feature to make it visually palatab
            channel_image -= channel_image.mean()
            channel_image /= channel_image.std()
            channel_image *= 64
            channel_image += 128
            channel_image = np.clip(channel_image, 0, 255).astype(
            display_grid[col * size : (col + 1) * size,
                         row * size : (row + 1) * size] = channel_

    # Display the grid
    scale = 1. / size
    plt.figure(figsize=(scale * display_grid.shape[1],
                        scale * display_grid.shape[0]))
    plt.title(layer_name)
    plt.grid(False)
    plt.imshow(display_grid, aspect='auto', cmap='viridis')

plt.show()
```

Experiment 4

```
model4 = models.Sequential()
model4.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides=
model4.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model4.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), strides
model4.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model4.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), strides
model4.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model4.add(layers.Flatten())
model4.add(layers.Dense(units=512, activation=tf.nn.relu))
model4.add(layers.Dense(units=10, activation=tf.nn.softmax))

model4.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history4 = model4.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model4.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history4.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
```

```
history_df.tail()

losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()

pred_classes = np.argmax(model4.predict(train_images_norm), axis=-

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

Experiment 5

```
model5 = models.Sequential()
model5.add(layers.Dense(units=64, activation=tf.nn.relu))
model5.add(layers.Dense(units=64, activation=tf.nn.relu))
model5.add(layers.Flatten())
model5.add(layers.Dense(units=10, activation=tf.nn.softmax))

model5.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history5 = model5.fit(train_images_norm
                      ,train_labels
                      ,epochs=20
                      ,batch_size=512
                      ,validation_data=(val_images_norm,val_labels)
```

```
                    ,callbacks=[
                    tf.keras.callbacks.EarlyStopping(monitor='val_
                    tf.keras.callbacks.ModelCheckpoint('/content/g
                                                        save_weights_only=False, m
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model5.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history5.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()

losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()

pred_classes = np.argmax(model5.predict(train_images_norm), axis=-

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

Experiment 6

```
model6 = models.Sequential()
```

```
model6.add(layers.Dense(units=64, activation=tf.nn.relu))
model6.add(layers.Dense(units=64, activation=tf.nn.relu))
model6.add(layers.Dense(units=64, activation=tf.nn.relu))
model6.add(layers.Flatten())
model6.add(layers.Dense(units=10, activation=tf.nn.softmax))

model6.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history6 = model6.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                    ,callbacks=[
                    tf.keras.callbacks.EarlyStopping(monitor='val_
                    tf.keras.callbacks.ModelCheckpoint('/content/g
                                        save_weights_only=False, m
)

time2 = perf_counter() - time
print(time2)

loss, accuracy = model6.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history6.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()

losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
```

```
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()

pred_classes = np.argmax(model6.predict(train_images_norm), axis=-

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

Experiment 7

```
model7 = models.Sequential()
model7.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides=
model7.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model7.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), strides
model7.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model7.add(layers.Flatten())
model7.add(layers.Dense(units=512, activation=tf.nn.relu))
model7.add(layers.Dense(units=10, activation=tf.nn.softmax))

model7.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history7 = model7.fit(train_images_norm
                      ,train_labels
                      ,epochs=20
                      ,batch_size=512
                      ,validation_data=(val_images_norm,val_labels)
                      ,callbacks=[
                      tf.keras.callbacks.EarlyStopping(monitor='val_
                      tf.keras.callbacks.ModelCheckpoint('/content/g
                                                save_weights_only=False, m
                     )
```

```
time2 = perf_counter() - time
print(time2)

loss, accuracy = model7.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history7.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()

losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()

pred_classes = np.argmax(model7.predict(train_images_norm), axis=-

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx

pred_classestest = np.argmax(model7.predict(test_images_norm), axi
conf_mx = tf.math.confusion_matrix(test_labels, pred_classestest)
conf_mx

from sklearn.metrics import precision_score
precision_score(test_labels, pred_classestest, average='micro')

precision_score(train_labels, pred_classes,  average='micro')
```

Experiment8

```
model8 = models.Sequential()
model8.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides=
model8.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model8.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), strides
model8.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model8.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), strides
model8.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model8.add(layers.Flatten())
model8.add(layers.Dense(units=512, activation=tf.nn.relu))
model8.add(layers.Dense(units=10, activation=tf.nn.softmax))

model8.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history8 = model8.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                    ,callbacks=[
                    tf.keras.callbacks.EarlyStopping(monitor='val_
                    tf.keras.callbacks.ModelCheckpoint('/content/g
                                        save_weights_only=False, m
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model8.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history8.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()
```

```
losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()
```

```
pred_classes = np.argmax(model8.predict(train_images_norm), axis=-
```

```
conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

```
pred_classestest = np.argmax(model8.predict(test_images_norm), axi
```

```
conf_mx = tf.math.confusion_matrix(test_labels, pred_classestest)
conf_mx
```

```
precision_score(test_labels, pred_classestest, average='micro')
```

```
precision_score(train_labels, pred_classes,  average='micro')
```

Experiment 9

```
model9 = models.Sequential()
model9.add(layers.Dense(units=128, kernel_regularizer='l1_l2', act
model9.add(layers.Dense(units=128, kernel_regularizer='l1_l2', act
model9.add(layers.Flatten())
model9.add(layers.Dense(units=10, kernel_regularizer='l1_l2', acti
```

```
model9.compile(optimizer='adam',
```

```
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history9 = model9.fit(train_images_norm
                     ,train_labels
                     ,epochs=20
                     ,batch_size=512
                     ,validation_data=(val_images_norm,val_labels)
                    )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model9.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history9.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()

losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()

pred_classes = np.argmax(model9.predict(train_images_norm), axis=-

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

Experiment 10

```
model10 = models.Sequential()
model10.add(layers.Dense(units=128, kernel_regularizer='l1_l2', ac
model10.add(layers.Dense(units=128, kernel_regularizer='l1_l2', ac
model10.add(layers.Dense(units=128, kernel_regularizer='l1_l2', ac
model10.add(layers.Flatten())
model10.add(layers.Dense(units=10, kernel_regularizer='l1_l2', act

model10.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history10 = model10.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model10.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history10.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()
```

Experiment 11

```
model11 = models.Sequential()
model11.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides
model11.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model11.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), stride
model11.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model11.add(layers.Flatten())
model11.add(layers.Dense(units=512, kernel_regularizer='l1_l2', ac
model11.add(layers.Dense(units=10, kernel_regularizer='l1_l2', act

model11.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history11 = model11.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model11.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)

history_dict = history11.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()

losses = history.history['loss']
accs = history.history['accuracy']
val_losses = history.history['val_loss']
val_accs = history.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
```

```
plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()

pred_classes = np.argmax(model10.predict(train_images_norm), axis=

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx
```

EXPERIMENT 12

```
model12 = models.Sequential()
model12.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides
model12.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model12.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), stride
model12.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model12.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), stride
model12.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model12.add(layers.Flatten())
model12.add(layers.Dense(units=512, kernel_regularizer='l1_l2', ac
model12.add(layers.Dense(units=10, kernel_regularizer='l1_l2', act

model12.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history12 = model12.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                   )

time2 = perf_counter() - time
print(time2)
```

```
print(time2)

loss, accuracy = model12.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)
```

Experiment 13

```
model13 = models.Sequential()
model13.add(layers.Dense(units=512, activation=tf.nn.relu))
model13.add(layers.Dense(units=512, activation=tf.nn.relu))
model13.add(Dropout(0.2))
model13.add(layers.Flatten())
model13.add(layers.Dense(units=10, activation=tf.nn.softmax))

model13.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history13 = model13.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model13.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)
```

Experiment 14

```
model14 = models.Sequential()
model14.add(layers.Dense(units=512, activation=tf.nn.relu))
model14.add(layers.Dense(units=512, activation=tf.nn.relu))
model14.add(Dropout(0.2))
```

```
model14.add(layers.Dense(units=512, activation=tf.nn.relu))
model14.add(layers.Flatten())
model14.add(layers.Dense(units=10, activation=tf.nn.softmax))

model14.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

time = perf_counter()
history14 = model14.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model14.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)
```

Experiment 15

```
model15 = models.Sequential()
model15.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides
model15.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model15.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), stride
model15.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model15.add(Dropout(0.2))
model15.add(layers.Flatten())
model15.add(layers.Dense(units=512,activation=tf.nn.relu))
model15.add(layers.Dense(units=10, activation=tf.nn.softmax))

model15.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])
```

```
time = perf_counter()
history15 = model15.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
                   )

time2 = perf_counter() - time
print(time2)

loss, accuracy = model15.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)
```

Experiment 16

```
model16 = models.Sequential()
model16.add(layers.Conv2D(filters=512, kernel_size=(3, 3), strides
model16.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model16.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), stride
model16.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model16.add(Dropout(0.2))
model16.add(layers.Conv2D(filters=1024, kernel_size=(3, 3), stride
model16.add(layers.MaxPooling2D(pool_size=(2, 2),strides=2))
model16.add(layers.Flatten())
model16.add(layers.Dense(units=512,activation=tf.nn.relu))
model16.add(layers.Dense(units=10,activation=tf.nn.softmax))

model16.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(f
              metrics=['accuracy'])

#time = perf_counter()
history16 = model16.fit(train_images_norm
                    ,train_labels
                    ,epochs=20
                    ,batch_size=512
                    ,validation_data=(val_images_norm,val_labels)
```

```
                    )
```

```
Epoch 1/20
79/79 [==============================] - 20s 223ms/step - loss: 2.1114 - accuracy: 0.2166
Epoch 2/20
79/79 [==============================] - 16s 208ms/step - loss: 1.4808 - accuracy: 0.4637
Epoch 3/20
79/79 [==============================] - 16s 207ms/step - loss: 1.2354 - accuracy: 0.5608
Epoch 4/20
79/79 [==============================] - 16s 207ms/step - loss: 1.0325 - accuracy: 0.6409
Epoch 5/20
79/79 [==============================] - 16s 207ms/step - loss: 0.9452 - accuracy: 0.6732
Epoch 6/20
79/79 [==============================] - 16s 207ms/step - loss: 0.8265 - accuracy: 0.7150
Epoch 7/20
79/79 [==============================] - 16s 207ms/step - loss: 0.7142 - accuracy: 0.7540
Epoch 8/20
79/79 [==============================] - 16s 207ms/step - loss: 0.6423 - accuracy: 0.7833
Epoch 9/20
79/79 [==============================] - 16s 207ms/step - loss: 0.5634 - accuracy: 0.8048
Epoch 10/20
79/79 [==============================] - 16s 207ms/step - loss: 0.5239 - accuracy: 0.8192
Epoch 11/20
79/79 [==============================] - 16s 207ms/step - loss: 0.4398 - accuracy: 0.8473
Epoch 12/20
79/79 [==============================] - 16s 207ms/step - loss: 0.3638 - accuracy: 0.8720
Epoch 13/20
79/79 [==============================] - 16s 207ms/step - loss: 0.2992 - accuracy: 0.8969
Epoch 14/20
79/79 [==============================] - 16s 207ms/step - loss: 0.2555 - accuracy: 0.9138
Epoch 15/20
79/79 [==============================] - 16s 207ms/step - loss: 0.2121 - accuracy: 0.9274
Epoch 16/20
79/79 [==============================] - 16s 207ms/step - loss: 0.1477 - accuracy: 0.9496
Epoch 17/20
79/79 [==============================] - 16s 207ms/step - loss: 0.1325 - accuracy: 0.9556
Epoch 18/20
79/79 [==============================] - 16s 207ms/step - loss: 0.1212 - accuracy: 0.9582
Epoch 19/20
79/79 [==============================] - 16s 207ms/step - loss: 0.1375 - accuracy: 0.9532
Epoch 20/20
79/79 [==============================] - 16s 207ms/step - loss: 0.0858 - accuracy: 0.9717
```

```
time2 = perf_counter() - time
print(time2)
```

```
loss, accuracy = model16.evaluate(test_images_norm, test_labels)
print('test set accuracy: ', accuracy * 100)
```

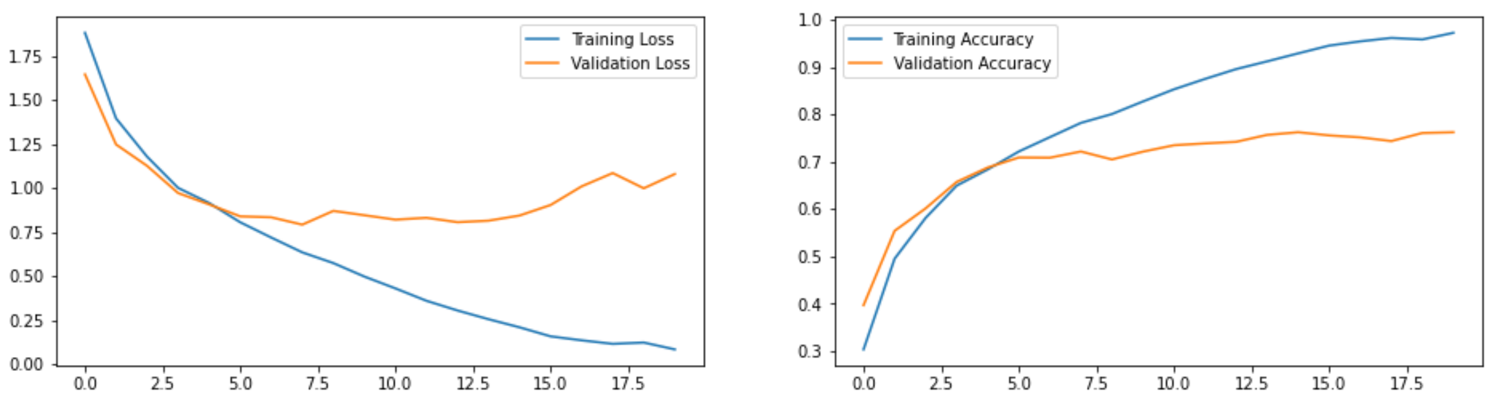
```
313/313 [==============================] - 3s 8ms/step - loss: 1.1476 - accuracy: 0.7483
test set accuracy:  74.83000159263611
```

```
history_dict = history16.history
history_dict.keys()
history_df=pd.DataFrame(history_dict)
history_df.tail()
```

	loss	accuracy	val_loss	val_accuracy
15	0.157892	0.945200	0.904099	0.7555
16	0.135336	0.954100	1.010051	0.7514
17	0.115368	0.961325	1.085448	0.7435
18	0.122782	0.958125	0.998898	0.7606
19	0.083785	0.972150	1.079883	0.7621

```
losses = history16.history['loss']
accs = history16.history['accuracy']
val_losses = history16.history['val_loss']
val_accs = history16.history['val_accuracy']
epochs = len(losses)

plt.figure(figsize=(16, 4))
for i, metrics in enumerate(zip([losses, accs], [val_losses, val_a
    plt.subplot(1, 2, i + 1)
    plt.plot(range(epochs), metrics[0], label='Training {}'.format
    plt.plot(range(epochs), metrics[1], label='Validation {}'.form
    plt.legend()
plt.show()
```




```
pred_classes = np.argmax(model16.predict(train_images_norm), axis=

conf_mx = tf.math.confusion_matrix(train_labels, pred_classes)
conf_mx

pred_classestest = np.argmax(model16.predict(test_images_norm), ax
```

```
conf_mx = tf.math.confusion_matrix(test_labels, pred_classestest)
conf_mx

precision_score(train_labels, pred_classes,  average='micro')

precision_score(test_labels, pred_classestest,  average='micro')

layer_outputs = [layer.output for layer in model16.layers]
activation_model = models.Model(inputs=model16.input, outputs=laye
layer_outputs

# Get the outputs of all the hidden nodes for each of the 60000 tr
activations = activation_model.predict(train_images_norm[0:1000])
hidden_layer_activation = activations[8]
output_layer_activations = activations[9]
hidden_layer_activation.shape   #  each of the 128 hidden nodes ha

#Get the dataframe of all the node values
activation_data = {'pred_class':pred_classes[0:1000]}
for k in range(0,512):
    activation_data[f"act_val_{k}"] = hidden_layer_activation[:,k]

activation_df = pd.DataFrame(activation_data)
activation_df.head()

# Separating out the features
features = [*activation_data][1:] # ['act_val_0', 'act_val_1',...]
x = activation_df.loc[:, features].values

pca = PCA(n_components=3)
principalComponents = pca.fit_transform(x)
principalDf = pd.DataFrame(data = principalComponents
             , columns = ['pca-one', 'pca-two', 'pca-three'])
principalDf.head()

pca.explained_variance_ratio_



activation_pca_df = pd.concat([principalDf, activation_df[['pred_c
activation_pca_df.head()
```

```
N=10000
activation_df_subset = activation_df.iloc[:N].copy()
activation_df_subset.shape

data_subset = activation_df_subset[features].values
data_subset.shape

from sklearn.manifold import TSNE

tsne = TSNE(n_components=2, verbose=1, perplexity=40, n_iter=300)
tsne_results = tsne.fit_transform(data_subset)


activation_df_subset['tsne-2d-one'] = tsne_results[:,0]
activation_df_subset['tsne-2d-two'] = tsne_results[:,1]

plt.figure(figsize=(16,10))
sns.scatterplot(
    x="tsne-2d-one", y="tsne-2d-two",
    hue="pred_class",
    palette=sns.color_palette("hls", 10),
    data=activation_df_subset,
    legend="full",
    alpha=0.3
)
```