

Bharati Vidyapeeth's College of Engineering for Women, Pune-43

Department of Information Technology

414447: Laboratory Practice IV (2019 pattern)

1. Study of Deep learning Packages: Tensorflow, Keras, Theano and PyTorch.
2. Implementing Feedforward neural networks with Keras and TensorFlow
 - a. Import the necessary packages
 - b. Load the training and testing data (MNIST/CIFAR10)
 - c. Define the network architecture using Keras
 - d. Train the model using SGD
 - e. Evaluate the network
 - f. Plot the training loss and accuracy
3. Build the Image classification model by dividing the model into following 4 stages:
 - a. Loading and preprocessing the image data
 - a. Defining the model's architecture
 - b. Training the model
 - c. Estimating the model's performance
4. Use Autoencoder to implement anomaly detection. Build the model by using:
 - a. Import required libraries
 - b. Upload / access the dataset
 - c. Encoder converts it into latent representation
 - d. Decoder networks convert it back to the original input
 - e. Compile the models with Optimizer, Loss, and Evaluation Metrics
5. Implement the Continuous Bag of Words (CBOW) Model. Stages can be:
 - a. Data preparation
 - b. Generate training data
 - c. Train model
 - d. Output
6. Object detection using Transfer Learning of CNN architectures.
 - a. Load in a pre-trained CNN model trained on a large dataset

- b. Freeze parameters (weights) in model's lower convolutional layers
- c. Add custom classifier with several layers of trainable parameters to model
- d. Train classifier layers on training data available for task
- e. Fine-tune hyper parameters and unfreeze more layers as needed

Assignment No.1

Title : Study of Deep learning Packages: Tensorflow, Keras, Theano and PyTorch.
Document the distinct features and functionality of the packages.

Aim: To Study and installation of following Deep learning Packages :

- i. Tensor Flow
- ii. Keras
- iii. Theano
- iv . PyTorch

Theory :

- 1) What is Deep Learning?
- 2) What are various packages in python for supporting Machine Learning libraries and which are mainly used for Deep Learning?
- 3) Compare Tensorflow / Keras/Theano and PyTorch on the following points (make a table) :
 - i. Available Functionality
 - ii. GUI status
 - iii. Versions
 - iv. Features
 - v. Compatibility with other environments.
 - vi. Specific Application domains.
- 4) Enlist the Models Datasets and pretrained models, Libraries and Extensions , Tools related to Tensorflow also discuss any two case studies like (PayPal, Intel, Etc.) related to TensorFlow. [Ref:<https://www.tensorflow.org/about>]
- 5) Explain the Keras Ecosystem.(kerastuner, kerasNLP, kerasCV, Autokeras and Model Optimization.) Also explain the following concepts related to keras : 1. Developing sequential Model 2. Training and validation using the inbuilt functions 3. Parameter Optimization. [Ref: <https://keras.io/>]
- 6) Explain simple Theano program.
- 7) Explain PyTorch Tensors . And also explain Uber's Pyro, Tesla Autopilot.[<https://pytorch.org/>]

Steps/ Algorithm :

Installation of Tensorflow On Ubuntu:

1. Install the Python Development Environment:

You need to download Python, the PIP package, and a virtual environment. If these packages are already installed, you can skip this step.

You can download and install what is needed by visiting the following links:

<https://www.python.org/>

<https://pip.pypa.io/en/stable/installing/>

<https://docs.python.org/3/library/venv.html>

To install these packages, run the following commands in the terminal:

```
sudo apt update
```

```
sudo apt install python3-dev python3-pip python3-venv
```

2. Create a Virtual Environment

Navigate to the directory where you want to store your Python 3.0 virtual environment. It can be in your home directory, or any other directory where your user can read and write permissions.

```
mkdir tensorflow_files
```

```
cd tensorflow_files
```

Now, you are inside the directory. Run the following command to create a virtual environment:

```
python3 -m venv virtualenv
```

The command above creates a directory named virtualenv. It contains a copy of the Python binary, the PIP package manager, the standard Python library, and other supporting files.

3. Activate the Virtual Environment

```
source virtualenv/bin/activate
```

Once the environment is activated, the virtual environment's bin directory will be added to the beginning of the \$PATH variable. Your shell's prompt will alter, and it will show the name of the virtual environment you are currently using, i.e. virtualenv.

4. Update PIP

```
pip install --upgrade pip
```

5. Install TensorFlow

The virtual environment is activated, and it's up and running. Now, it's time to install the TensorFlow package.

```
pip install -- upgrade TensorFlow
```

Installation of Keras on Ubuntu :

Prerequisite : Python version 3.5 or above.

STEP 1: Install and Update Python3 and Pip

Skip this step if you already have Python3 and Pip on your machine.

```
sudo apt install python3 python3.pip
```

```
sudo pip3 install --upgrade pip
```

STEP 2: Upgrade Setuptools

```
pip3 install --upgrade setuptools
```

STEP 3: Install TensorFlow

```
pip3 install tensorflow
```

Verify the installation was successful by checking the software package information:

```
pip3 show tensorflow
```

STEP 4: Install Keras

```
pip3 install keras
```

Verify the installation by displaying the package information:

```
pip3 show keras  
[https://phoenixnap.com/kb/how-to-install-keras-on-linux]
```

Installation of Theano on Ubuntu:

Step 1: First of all, we will install Python3 on our Linux Machine. Use the following command in the terminal to install Python3.

```
sudo apt-get install python3
```

Step 2: Now, install the pip module

```
sudo apt install python3-pip
```

Step 3: Now, install the Theano

Verifying Theano package Installation on Linux using PIP

```
python3 -m pip show theano
```

Installation of PyTorch

First, check if you are using python's latest version or not. Because PyGame requires python 3.7 or a higher version

```
python3 --version
```

```
pip3 --version
```

```
pip3 install torch==1.8.1+cpu torchvision==0.9.1+cpu  
torchaudio==0.8.1 -f
```

https://download.pytorch.org/whl/torch_stable.html

[Ref : <https://www.geeksforgeeks.org/install-pytorch-on-linux/>]

Python Libraries and functions required

1. Tensorflow,keras

numpy : NumPy is a Python library used for working with arrays. It also has functions for working in the domain of linear algebra, fourier transform, and matrices. NumPy stands for Numerical Python.

To import numpy use

```
import numpy as np
```

pandas: pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

To import pandas use

```
import pandas as pd
```

sklearn : Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistent interface in Python. This library, which is largely written in Python, is built upon NumPy, SciPy and Matplotlib.

For importing train_test_split use

```
from sklearn.model_selection import train_test_split
```

2. For Theano Requirements:

- Python3

- Python3-pip
- NumPy
- SciPy
- BLAS

Sample Code with comments

1. Tensorflow Test program:

```
import tensorflow as tf

print(tf.__version__)

2.1.0

print(tf.reduce_sum(tf.random.normal([1000, 1000])))

tf.Tensor(-505.04108, shape=(), dtype=float32)
```

2. Keras Test Program:

```
from tensorflow import keras
from keras import datasets
```

```
# Load MNIST data
# (train_images, train_labels), (test_images, test_labels) = datasets.mnist.load_data()

# Check the dataset loaded
#train_images.shape, test_images.shape
```

3. Theano test program

```
# Python program showing addition of two scalars

# Addition of two scalars
import numpy
import theano.tensor as T
from theano import function

# Declaring two variables
x = T.dscalar('x')
y = T.dscalar('y')
```

```
# Summing up the two numbers  
z = x + y
```

Converting it to a callable object so that it takes matrix as parameters

```
f = function([x, y], z)  
f(5, 7)
```

4. Test program for PyTorch

```
#The usual imports  
import torch  
import torch.nn as nn
```

```
#print out the pytorch version used  
print(torch.__version__)
```

Output of Code:

Note: Run the code and attach your output of the code here.

Conclusion :

Tensorflow, PyTorch, Keras and Theano all these packages are installed and ready for Deep learning applications. As per application domain and dataset we can choose the appropriate package and build the required type of Neural Network.

Assignment No. 2

Problem Statement

To Implement Feed forward Neural Network with keras and tensorflow.

1. Import the necessary packages
2. Load the training and testing data(MNIST)
3. Define the network architecture using keras
4. Train the model using SGD
5. Evaluate the network
6. Plot the training loss and accuracy

Solution Expected

Implement and train a feed-forward neural network (also known as an "MLP" for "multi-layer perceptron") on a dataset called MNIST and improve model generalization by achieving increased accuracy and decreased loss where model gains good confidence with the prediction.

Objectives to be achieved

1. Understand how to use Tensorflow Eager and Keras Layers to build a neural network architecture.
2. Understand how a model is trained and evaluated.
3. Identify digits from images.
4. Our main goal is to train a neural network (using Keras) to obtain $> 90\%$ accuracy on MNIST dataset.
5. Research at least 1 technique that can be used to improve model generalization.

Methodology to be used

- Deep Learning
- Feed Forward Neural Network

Justification with Theory/Literature

Deep learning has revolutionized the world of machine learning as more and more ML practitioners have adopted deep learning networks to solve real-world problems. Compared to the more traditional ML models, deep learning networks have been shown superior performance for many applications.

The first step toward using deep learning networks is to understand the working of a simple feedforward neural network. We get started with how we can build our first neural network model using **Keras** running on top of the **Tensorflow** library.

TensorFlow is an open-source platform for machine learning. Keras is the high-level application programming interface (API) of TensorFlow. Using Keras, we can rapidly develop a prototype system and test it out. This is the first in a three-part series on using TensorFlow for supervised classification tasks.

A Conceptual Diagram of the Neural Network

We'll build a supervised classification model that learns to identify digits from images. We'll use the well-known MNIST dataset to train and test our model. The MNIST dataset consists of 28-by-28 images of handwritten digits along with their corresponding labels.

We'll construct a neural network with one hidden layer to classify each digit image into its corresponding label. The figure below shows a conceptual diagram of the neural network we are about to build. The output layer consists of 10 units, where each unit corresponds to a different digit. Each unit computes a value that can be interpreted as the confidence value of its respective digit. The final classification is the digit with the maximum confidence value.

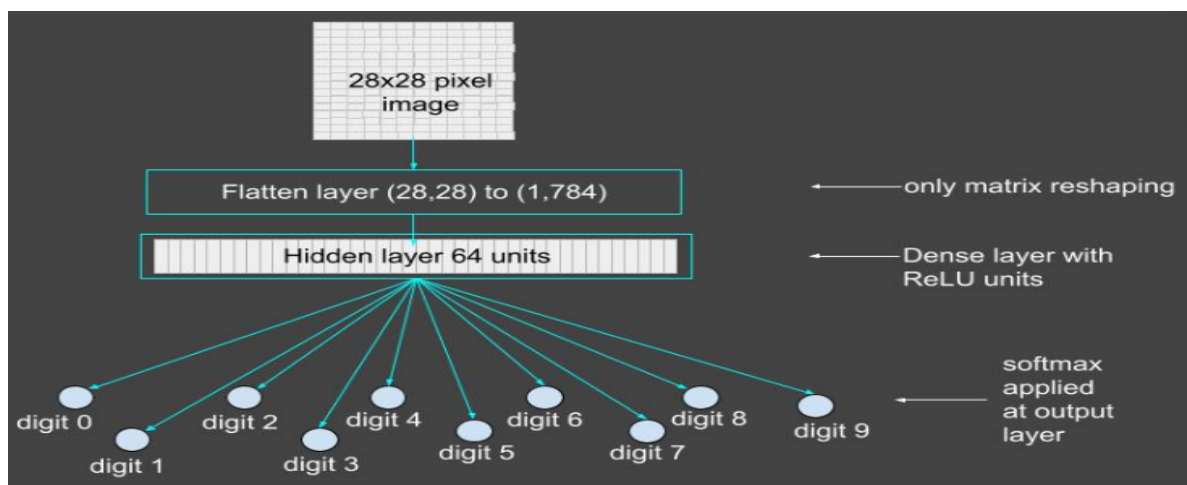


Figure 1: Conceptual diagram of the neural network.

Each output unit corresponds to a digit and has its confidence value. The final classification is the digit with the maximum confidence value.

TensorFlow and Keras Libraries

If Keras and TensorFlow are not installed on your system, you can easily do so using pip or conda depending upon your Python environment.

```
pip install tensorflow
```

In the context of ML, a tensor is a multidimensional array, which in its simplest form is a scalar. Vectors and matrices are special cases of tensors. In TensorFlow, a tensor is a data structure. It is a multidimensional array composed of elements of the same type. Tensors are used to encapsulate all inputs and outputs to a deep learning network. The training dataset and each test example has to be cast as a tensor. All operations within the layers of the network are also performed on tensors.

Layers in TensorFlow?

You can build a fully connected feedforward neural network by stacking layers sequentially so that the output of one layer becomes the input to the next. In TensorFlow, layers are callable objects, which take tensors as input and generate outputs that are also tensors. Layers can contain weights and biases, which are both tuned during the training phase. We'll create a simple neural network from two layers:

1. Flatten layer
2. Dense layer

The Flatten Layer:

This layer flattens an input tensor without changing its values. Given a tensor of rank n , the Flatten layer reshapes it to a tensor of rank 2. The number of elements on the first axis remains unchanged. The elements of the input tensor's remaining axes are stacked together to form a single axis. We need this layer to create a vectorized version of each image for further input to the next layer.

The Dense Layer:

The dense layer is the fully connected, feedforward layer of a neural network. It computes the weighted sum of the inputs, adds a bias, and passes the output through an activation function. We are using the ReLU activation function for this example. This function does not change any value greater than 0. The rest of the values are all set to 0.

The computations of this layer for the parameters shown in the code above are all illustrated in the figure below.

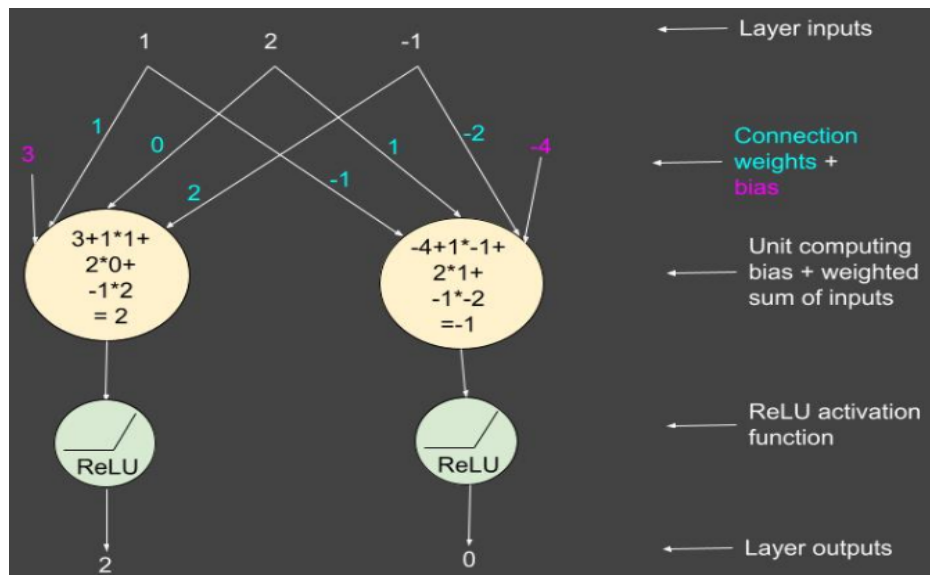


Figure 2: Hidden layer computations.

Creating a Model in TensorFlow

We are now ready to create a model of a simple neural network with one hidden layer. The simplest method is to use `Sequential()` with a list of all the layers you want to stack together. The code below creates a model and prints its summary. Note the use of `relu` as the activation function in the first dense layer and a `softmax` in the output layer. The `softmax` function normalizes all outputs to sum to 1 and ensures that they are in the range `[0, 1]`.

```
model = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(64, activation='relu'),
    tf.keras.layers.Dense(10, activation='softmax')
])
```

Compile the Model

Next we compile the model. Here, we specify the optimizer and loss function of our model. The optimization algorithm determines how the connection weights are updated at each training step with respect to the given loss function. Because we have a multiclass classification problem, the loss function we'll use is `categorical_crossentropy`, coupled with the adam optimizer. You can experiment with other optimizers too. The value of the metrics argument sets the parameter to monitor and record during the training phase.

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

Train the Neural Network

Now that the model is ready, it's time to train it. We'll load the dataset, train the model, and view the training process. Note that the outputs shown here will vary with every run of the program because of the stochastic nature of the algorithms involved.

Load the Dataset

The following code loads the training set and the test set of the MNIST data. It also prints the statistics of both sets. Because our model has 10 outputs, one for each digit, we need to convert the absolute image labels to categorical ones. The `utils` module in the Keras library provides the method `to_categorical()` for this conversion.

Train the Model

The `fit()` method of the model object trains the neural network. If you want to use a validation set during training, all you have to do is define the percentage of validation examples to be taken from the training set. The splitting of the training set into a train and validation set is automatically taken care of by the `fit()` method.

In the code below, the `fit()` method is called in 10 epochs.

View the Training Process

The `fit()` method returns a history object with detailed information regarding model training. The `history` attribute is a dictionary object.

To view the learning process, we can plot the accuracy and loss corresponding to different epochs for both the training and validation sets. The following code creates two graphs for these two metrics.

The `predict()` method

If you want to see the output of the network for one or more train/test examples, use the `predict()` method. The following example code prints the values of the output layer when the first test image is used as an input. It is a 10-dimensional vector of confidence values corresponding to each digit. The final classification of the image is the `argmax()` of this vector.

The `evaluate()` method

The `evaluate()` method computes the overall accuracy of the model on the dataset passed as an argument. The code snippet below prints the classification accuracy on both the training set and the test set.

Code Snippets

▼ Importing Libraries

```
✓ [1] #import necessary libraries  
2s import tensorflow as tf  
from tensorflow import keras
```

```
✓ [2] import pandas as pd  
0s import numpy as np  
import matplotlib.pyplot as plt  
import random  
%matplotlib inline
```

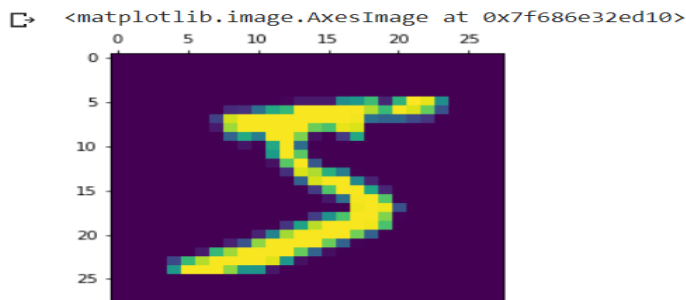
▸ Loading & Preparing the data

MNIST stands for Modified National Institute of Standards and Technology dataset. It is a dataset of 70,000 handwritten images. Each image is of 28*28 pixel i.e about 784 features. Each feature represent only one pixel intensity i.e from 0(white) to 255(black). This dataset is further divided into 60000 training and 10000 testing images.

```
#import dataset and split into train and test
mnist = tf.keras.datasets.mnist
(x_train,y_train),(x_test,y_test) = mnist.load_data()

Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/mnist.npz
11493376/11490434 [=====] - 0s 0us/step
11501568/11490434 [=====] - 0s 0us/step
```

```
#to see how first image looks
plt.matshow(x_train[0])
```



normalizing the images by scaling the pixel intensities to the range 0 to 1

```
[ ] #normalizing the images by scaling the pixel intensities to the range 0 to 1
x_train = x_train/255
x_test = x_test/255
```

▶ `x_train[0]`

```

↳
[0. , 0. , 0. ],
0. , 0. , 0. , 0. , 0. ,
0. , 0. , 0.19215686, 0.93333333, 0.99215686,
0.99215686, 0.99215686, 0.99215686, 0.99215686, 0.99215686,
0.99215686, 0.99215686, 0.98431373, 0.36470588, 0.32156863,
0.32156863, 0.21960784, 0.15294118, 0. ,
0. , 0. , 0. ],
[0. , 0. , 0. , 0. , 0. , 0. ,
0. , 0. , 0.07058824, 0.85882353, 0.99215686,
0.99215686, 0.99215686, 0.99215686, 0.99215686, 0.77647059,
0.71372549, 0.96862745, 0.94509804, 0. , 0. ,
0. , 0. , 0. , 0. , 0. , 0. ,
0. , 0. , 0. ],
[0. , 0. , 0. , 0. , 0. , 0. ,
0. , 0. , 0. , 0.31372549, 0.61176471,
0.41960784, 0.99215686, 0.99215686, 0.80392157, 0.04313725,
0. , 0.16862745, 0.60392157, 0. , 0. ]

```

✓ 0s completed at 2:22 PM

▼ Creating the model

The Relu function is one of the most popular activation function. It stands for "Rectified Linear Unit". Mathematically this function is defined as $y = \max(0, x)$. The relu function returns 0 if the input is negative and linear if the input is positive.

The softmax function is another activation function. It changes input values into values that reach from 0 to 1.

```
[ ] model = keras.Sequential([
    keras.layers.Flatten(input_shape=(28,28)),
    keras.layers.Dense(128,activation='relu'),
    keras.layers.Dense(10,activation='softmax')
])
```

```
[ ] model.summary()
```

Model: "sequential_2"

Layer (type)	Output Shape	Param #
flatten_2 (Flatten)	(None, 784)	0
dense_4 (Dense)	(None, 128)	100480
dense_5 (Dense)	(None, 10)	1290

```
=====
Total params: 101,770
Trainable params: 101,770
Non-trainable params: 0
=====
```

▼ Compile the model

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


▼ Train the model

```
[ ] history=model.fit(x_train,y_train,validation_data=(x_test,y_test),epochs=10)
```

```
Epoch 1/10
1875/1875 [=====] - 5s 3ms/step - loss: 0.6254 - accuracy: 0.8444 - val_loss: 0.3584 - val_accuracy: 0.9009
Epoch 2/10
1875/1875 [=====] - 5s 2ms/step - loss: 0.3366 - accuracy: 0.9053 - val_loss: 0.3012 - val_accuracy: 0.9156
Epoch 3/10
1875/1875 [=====] - 5s 3ms/step - loss: 0.2889 - accuracy: 0.9188 - val_loss: 0.2641 - val_accuracy: 0.9252
Epoch 4/10
1875/1875 [=====] - 5s 2ms/step - loss: 0.2583 - accuracy: 0.9276 - val_loss: 0.2405 - val_accuracy: 0.9330
Epoch 5/10
1875/1875 [=====] - 5s 3ms/step - loss: 0.2352 - accuracy: 0.9345 - val_loss: 0.2215 - val_accuracy: 0.9384
Epoch 6/10
1875/1875 [=====] - 5s 2ms/step - loss: 0.2162 - accuracy: 0.9392 - val_loss: 0.2047 - val_accuracy: 0.9430
Epoch 7/10
1875/1875 [=====] - 5s 3ms/step - loss: 0.2004 - accuracy: 0.9437 - val_loss: 0.1907 - val_accuracy: 0.9455
Epoch 8/10
1875/1875 [=====] - 5s 3ms/step - loss: 0.1871 - accuracy: 0.9472 - val_loss: 0.1821 - val_accuracy: 0.9476
Epoch 9/10
1875/1875 [=====] - 5s 2ms/step - loss: 0.1755 - accuracy: 0.9505 - val_loss: 0.1709 - val_accuracy: 0.9501
Epoch 10/10
1875/1875 [=====] - 5s 2ms/step - loss: 0.1653 - accuracy: 0.9539 - val_loss: 0.1622 - val_accuracy: 0.9521
```

✓ 0s completed at 2:22 PM

▼ Evaluate the model

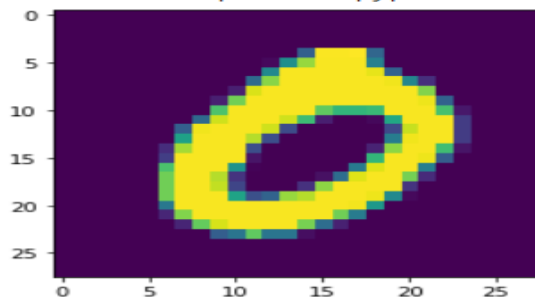
```
[ ] test_loss,test_acc=model.evaluate(x_test,y_test)
print("Loss=%.3f" %test_loss)
print("Accuracy=%.3f" %test_acc)
```

```
313/313 [=====] - 1s 2ms/step - loss: 0.1622 - accuracy: 0.9521
Loss=0.162
Accuracy=0.952
```

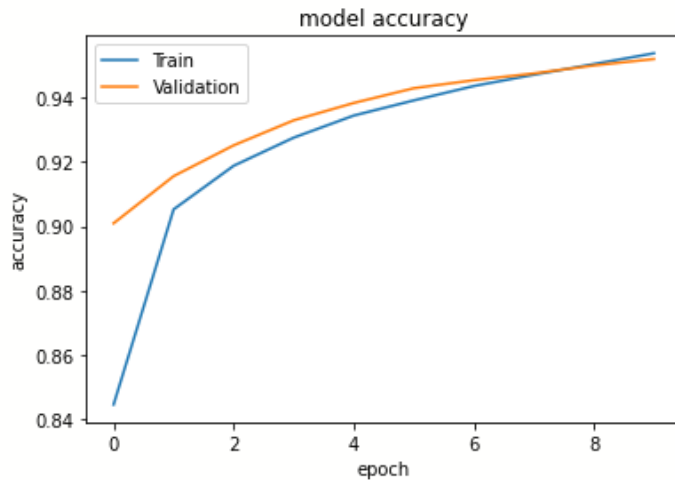
▼ Making prediction on new data

```
▶ n=random.randint(0,9999)
plt.imshow(x_test[n])
plt.show
```

```
↳ <function matplotlib.pyplot.show(*args, **kw)>
```



Plot graph for accuracy and loss



Confusion Matrix

```

test_predict = model.predict(x_test)
# Get the classification labels
test_predict_labels = np.argmax(test_predict, axis=1)
confusion_matrix = tf.math.confusion_matrix(labels=y_test, predictions=test_predict_labels)
print('Confusion matrix of the test set:\n', confusion_matrix)

```

Confusion matrix of the test set:

```

tf.Tensor(
[[ 965    0    1    1    0    4    6    1    2    0]
 [    0 1117    2    2    1    1    3    2    7    0]
 [    6    2  977   13    5    1    7    8   12    1]
 [    1    0    9  960    1    9    1   10   13    6]
 [    1    1    4    0  939    1    8    3    3   22]
 [    9    1    2   19    3  823   11    2   14    8]
 [    9    3    2    2   10    7  919    1    5    0]
 [    1   10   18    7    5    1    0  968    1   17]
 [    3    1    3   18    7    7    7    9  916    3]
 [    8    8    0   12   22    3    1   12    6  937]], shape=(10, 10), dtype=int32)

```

Conclusion

With above code we can see that, throughout the epochs, our model accuracy increases and loss decreases that is good since our model gains confidence with our prediction

This indicates the model is trained in a good way

1. The two loss(loss and val_loss) are decreasing and the accuracy (accuracy and val_accuracy) increasing.

2. The val_accuracy is the measure of how good the model is predicting so, it is observed that the model is well trained after 10 epochs

References

1. S. Arora and M. P. S. Bhatia, "Handwriting recognition using Deep Learning in Keras," *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 2018, pp. 142-145, doi: 10.1109/ICACCCN.2018.8748540.
2. <https://towardsdatascience.com/feed-forward-neural-networks-how-to-successfully-build-them-in-python-74503409d99a>
3. <https://pyimagesearch.com/2021/05/06/implementing-feedforward-neural-networks-with-keras-and-tensorflow/>
4. <https://exchange.scale.com/public/blogs/how-to-build-a-fully-connected-feedforward-neural-network-using-keras-and-tensorflow>
5. <https://www.kaggle.com/code/prashant111/mnist-deep-neural-network-with-keras/notebook>
6. <https://sanjayasubedi.com.np/deeplearning/tensorflow-2-first-neural-network-for-fashion-mnist/>

Assignment No. 3

Problem Statement:

Build the Image classification model by dividing the model into following 4 stages:

1. Loading and preprocessing the image data
2. Defining the model's architecture
3. Training the model
4. Estimating the model's performance

Objective:

1. To be able to apply deep learning algorithms to solve problems of moderate complexity
2. Understand how a model is trained and evaluated.
3. Classifying images from the image dataset.
4. Our main goal is to train a neural network (using Keras) to obtain > 90% accuracy on image dataset..
5. To apply the algorithms to a real-world problem, optimize the models learned and report on the expected accuracy that can be achieved by applying the models

Outcomes:

At the end of the assignment the students should able-

1. Learn and Use various Deep Learning tools and packages.
2. Build and train deep Neural Network models for use in various applications.
3. Apply Deep Learning techniques like CNN, RNN Auto encoders to solve real word Problems.
4. Evaluate the performance of the model build using Deep Learning.

Solution Expected:

Implement and train a Convolutional neural network (CNN) on an hand-written digits image dataset called MNIST and improve model generalization by achieving increased accuracy and decreased loss where model gains good confidence with the prediction.

Methodology to be used

Deep Learning

Convolutional Neural Network

Theory:

Deep Learning has been proved that its a very powerful tool due to its ability to handle huge amounts of data. The use of hidden layers exceeds traditional techniques, especially for pattern recognition. One of the most popular Deep Neural Networks is Convolutional Neural Networks (CNN).

Convolutional Neural Networks (CNNs)

A convolutional neural network (CNN) is a type of Artificial Neural Network (ANN) used in image recognition and processing which is specially designed for processing data (pixels). The goal of a CNN is to learn higher-order features in the data via convolutions. They are well suited to object recognition with images and consistently top image classification competitions. They can identify faces, individuals, street signs, platypuses, and many other aspects of visual data. CNNs overlap with text analysis via optical character recognition, but they are also useful when analyzing words⁶ as discrete textual units. They're also good at analyzing sound. The efficacy of CNNs in image recognition is one of the main reasons why the world recognizes the power of deep learning. CNNs are good at building position and (somewhat) rotation invariant features from raw image data. CNNs are powering major advances in machine vision, which has obvious applications for self-driving cars, robotics, drones, and treatments for the visually impaired. The structure of image data allows us to change the architecture of a neural network in a way that we can take advantage of this structure. With CNNs, we can arrange the neurons in a three-dimensional structure for which we have the following:

1. Width
2. Height
3. Depth

These attributes of the input match up to an image structure for which we have:

1. Image width in pixels
2. Image height in pixels
3. RGB channels as the depth

We can consider this structure to be a three-dimensional volume of neurons. A significant aspect to how CNNs evolved from previous feed-forward variants is how they achieved computational efficiency with new layer types.

CNN Architecture Overview

CNNs transform the input data from the input layer through all connected layers into a set of class scores given by the output layer. There are many variations of the CNN architecture, but they are based on the pattern of layers, as demonstrated in Figure 1 (below).

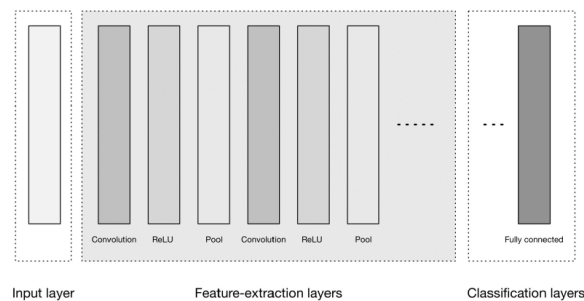


Figure 1. High-level general CNN architecture

Figure 1 depicts three major groups:

1. Input layer

- An image matrix (volume) of dimension **$(h \times w \times d)$**
- A filter **$(f_h \times f_w \times d)$**
- Outputs a volume dimension **$(h - f_h + 1) \times (w - f_w + 1) \times 1$**

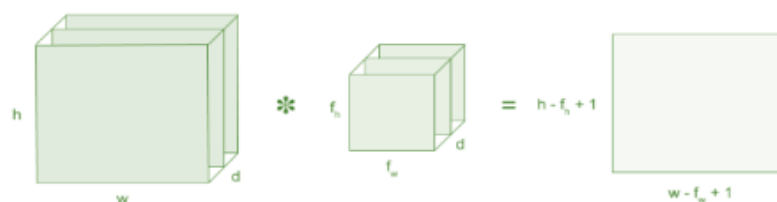


Figure 2: Image matrix multiplies kernel or filter matrix

The input layer accepts three-dimensional input generally in the form spatially of the size (width \times height) of the image and has a depth representing the color channels (generally three for RGB color channels) as shown in fig 2 above.

2. Feature-extraction (learning) layers

The feature-extraction layers have a general repeating pattern of the sequence as shown in fig 1 above of :

A. Convolution layer

We express the Rectified Linear Unit (ReLU) activation function as a layer in the diagram in fig 1. convolutional layers transform the input data by using a patch of locally connecting neurons from the previous layer

B. Pooling layer

These layers find a number of features in the images and progressively construct higher-order features. This corresponds directly to the ongoing theme in deep learning by which features are automatically learned as opposed to traditionally hand engineered.

3. Classification layers

Finally we have the classification layers in which we have one or more fully connected layers to take the higher-order features and produce class probabilities or scores. These layers are fully connected to all of the neurons in the previous layer, as their name implies. The output of these layers produces typically a two dimensional output of the dimensions $[b \times N]$, where b is the number of examples in the mini-batch and N is the number of classes we're interested in scoring.

Multilayer neural networks vs CNN

In traditional multilayer neural networks, the layers are fully connected and every neuron in a layer is connected to every neuron in the next layer whereas The neurons in the layers of a CNN are arranged in three dimensions to match the input volumes. Here, depth means the third dimension of the activation volume, not the number of layers, as in a multilayer neural network.

Evolution of the connections between layers

Another change is how we connect layers in a convolutional architecture. Neurons in a layer are connected to only a small region of neurons in the layer before it. CNNs retain a layer-oriented architecture, as in traditional multilayer networks, but have different types of layers. Each layer transforms the 3D input volume from the previous layer into a 3D output volume of neuron activations with some differentiable function that might or might not have parameters, as demonstrated in Figure 1.

Input Layers

Input layers are where we load and store the raw input data of the image for processing in the

network. This input data specifies the width, height, and number of channels. Typically, the number of channels is three, for the RGB values for each pixel.

Convolutional Layers

Convolutional layers are considered the core building blocks of CNN architectures. As Figure 2 illustrates, convolutional layers transform the input data by using a patch of locally connecting neurons from the previous layer. The layer will compute a dot product between the region of the neurons in the input layer and the weights to which they are locally connected in the output layer.

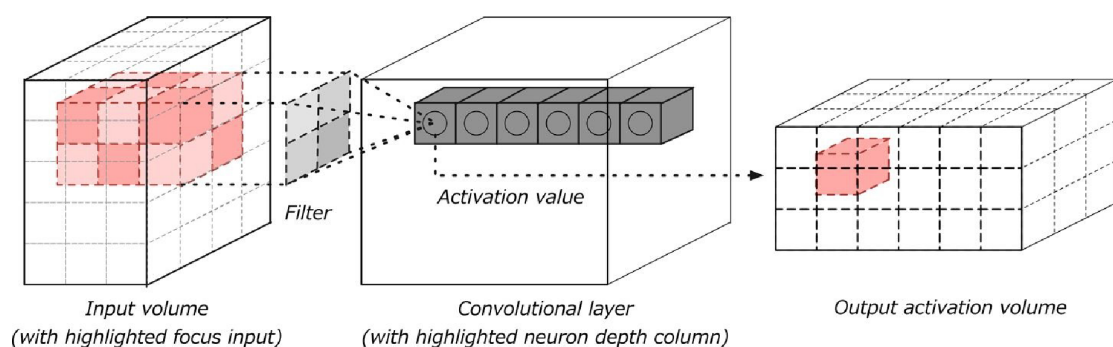


Figure 2. Convolution layer with input and output volumes

The resulting output generally has the same spatial dimensions (or smaller spatial dimensions) but sometimes increases the number of elements in the third dimension of the output (depth dimension).

TensorFlow is an open-source platform for machine learning. Keras is the high-level application programming interface (API) of TensorFlow. Using Keras, we can rapidly develop a prototype system and test it out. This is the first in a three-part series on using TensorFlow for supervised classification tasks.

STEPS: To implement Convolutional Neural Network for image classification

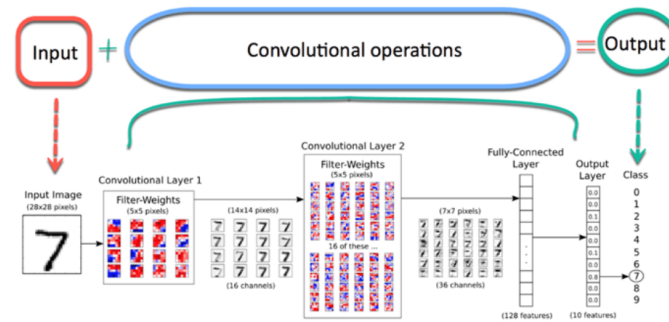


Figure 2. Image Classification model using CNN with python

CNN:

Now imagine there is an image of a bird, and you want to identify it whether it is really a bird or something else. The first thing you should do is feed the pixels of the image in the form of arrays to the input layer of the neural network (MLP networks used to classify such things). The hidden layers carry Feature Extraction by performing various calculations and operations. There are multiple hidden layers like the convolution, the ReLU, and the pooling layer that performs feature extraction from your image. So finally, there is a fully connected layer that you can see which identifies the exact object in the image. You can understand very easily from the following figure:

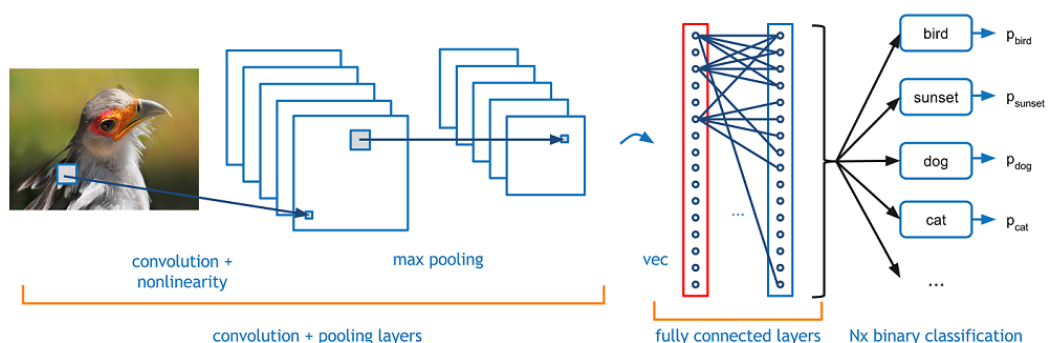


Figure 2. Convolution and pooling layers in CNN

Convolution:-

Convolution Operation involves matrix arithmetic operations and every image is represented in the form of an array of values(pixels).

Let us understand example:

$$a = [2,5,8,4,7,9]$$

$$b = [1,2,3]$$

In Convolution Operation, the arrays are multiplied one by one element-wise, and the product is

grouped or summed to create a new array that represents $\mathbf{a} * \mathbf{b}$.

The first three elements of matrix \mathbf{a} are now multiplied by the elements of matrix \mathbf{b} . The product is summed to get the result and stored in a new array of $\mathbf{a} * \mathbf{b}$.

This process remains continuous until the operation gets completed.

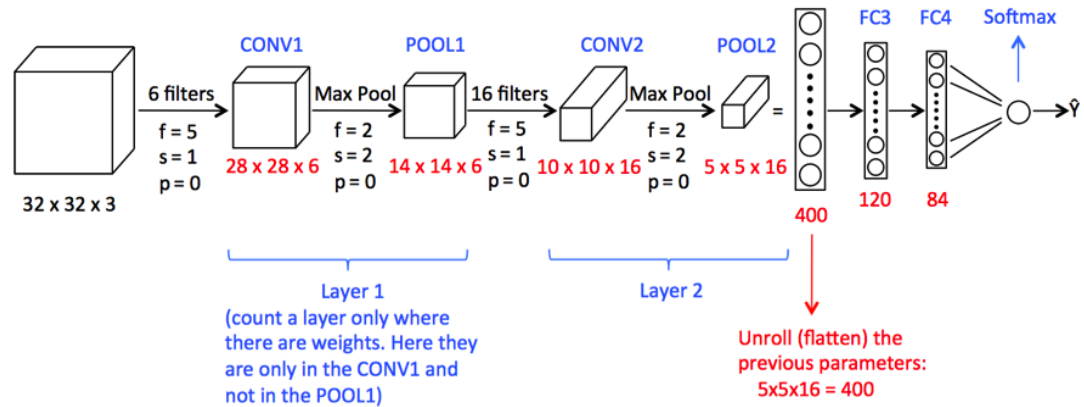


Figure 2. Sequence of Convolution and pooling layers in CNN

Pooling:

After the convolution, there is another operation called pooling. So, in the chain, convolution and pooling are applied sequentially on the data in the interest of extracting some features from the data. After the sequential convolutional and pooling layers, the data is flattened into a feed-forward neural network which is also called a Multi-Layer Perceptron.

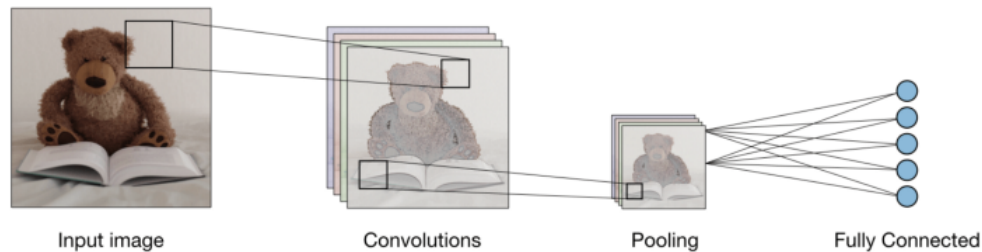


Fig: Data flattening into a feed-forward neural network

Thus, we have seen steps that are important for building CNN model.

Code Snippets

```
#Importing Libraries  
  
import numpy as np
```

```
import pandas as pd
import random
import tensorflow as tf
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score

from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Flatten, Conv2D, Dense,
MaxPooling2D
from tensorflow.keras.optimizers import SGD
from tensorflow.keras.utils import to_categorical
from tensorflow.keras.datasets import mnist
```

Loading and preprocessing the image data

Loading the MNIST dataset is very simple using Keras. The dataset is so popular that you can access it through `datasets.mnist` and use the `load_data()` function to get a train and a test set.

The dataset contains 28x28 images showing handwritten digits.

```
(X_train, y_train), (X_test, y_test) = mnist.load_data()
```

If you print the shape of any of the sets you will see more information about the samples.

For example, printing the shape of the train set will get you (60000, 28, 28):

60000: This is the number of samples in the set.

28: This is the height of each image.

28: This is the width of each image.

So we have 60,000 28x28 images.

```
print(X_train.shape)
(60000, 28, 28)
```

Anytime you are using a neural network, you should pay special attention to the range of

the input values you will be feeding it. In our case, each image is a matrix of 28x28 pixels.

Let's print the range of these values to understand what's the scale we are working with:

```
X_train[0].min(), X_train[0].max()  
(0, 255)
```

The pixel values in our images are between 0 (black) to 255 (white).

Neural networks have a much easier time converging when they work with values that don't vary a lot in scale. It's a common practice to scale every input to fit a small range like 0 to 1 or -1 to 1. We should do that here, and scale our pixels to a range that goes from 0 (black) to 1 (white).

Here is the formula to scale a value: $\text{scaled_value} = (\text{original_value} - \text{min}) / (\text{max} - \text{min})$. In our case, the minimum value is 0 and the maximum value is 255.

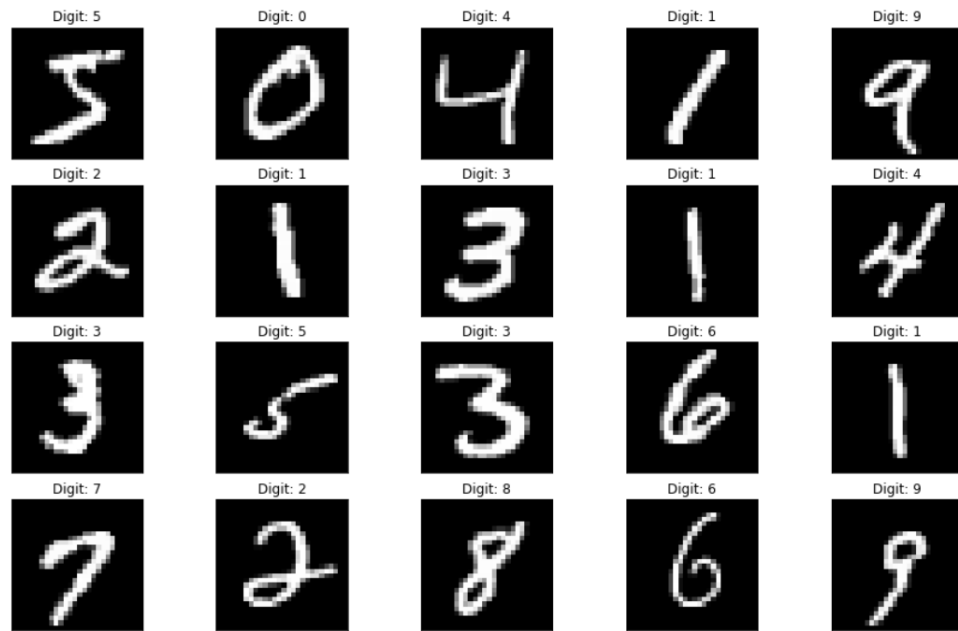
Let's use this to scale our train and test sets (notice that you can get rid of the 0.0 in the formula below but I'm leaving them there for clarity purposes):

```
X_train = (X_train - 0.0) / (255.0 - 0.0)  
X_test = (X_test - 0.0) / (255.0 - 0.0)  
X_train[0].min(), X_train[0].max()  
(0.0, 1.0)
```

We can now plot the first 20 images on the train set:

```
def plot_digit(image, digit, plt, i):  
    plt.subplot(4, 5, i + 1)  
    plt.imshow(image, cmap=plt.get_cmap('gray'))  
    plt.title(f"Digit: {digit}")  
    plt.xticks([])  
    plt.yticks([])
```

```
plt.figure(figsize=(16, 10))
for i in range(20):
    plot_digit(X_train[i], y_train[i], plt, i)
plt.show()
```



In Computer Vision, we usually use 4 dimensions to represent a set of images:

- The total number of images (we call this "batch size")
- The width of each image
- The height of each image
- The number of channels of each image

As you saw before, our train set has 3 dimensions only; we are missing the number of channels. We need to transform our data by adding that fourth dimension. Since these images are grayscale, that fourth dimension will be 1.

We can use numpy's `reshape()` function to reshape all of the data by adding that extra dimension.

```
X_train = X_train.reshape((X_train.shape + (1,)))
X_test = X_test.reshape((X_test.shape + (1,)))
```

Finally, let's take a look at the format of our target values (`y_train`). Let's print the first 20 samples in our train set:

```
y_train[0:20]  
array([5, 0, 4, 1, 9, 2, 1, 3, 1, 4, 3, 5, 3, 6, 1, 7, 2,  
      8, 6, 9],  
      dtype=uint8)
```

We are representing the target digits using integers (the digit 5 is represented with value 5, the digit 0 with value 0, etc.) This is important because it determines which loss function we should use to optimize our neural network.

We have two options:

1. Use integers for our target values (as they are now), and make sure we use the "Sparse Categorical Cross-Entropy" loss function.
2. One-hot encode the targets and use the "Categorical Cross-Entropy" loss function.

The easiest solution is to leave the targets as they are, so let's move on to creating the model.

Defining the model's architecture

There are several ways to create a model in Keras. In this example, we are going to use Keras's Sequential API because it's very simple to use.

Let's break down the definition of model below step by step:

First, we are going to define the first hidden layer of our network: A convolutional layer with 32 filters and a 3x3 kernel. This layer will use a ReLU activation function. The goal of this layer is to generate 32 different representations of an image, each one of 26x26. The 3x3 kernel will discard a pixel on each side of the original image and that's way we get 26x26 squares instead of 28x28.

Notice how we also need to define the input shape of the network as part of that first layer. Remember that our images are 28x28 with a single color channel, so that leads to the (28, 28, 1) shape.

Right after that first layer, we are going to do a 2x2 max pooling to downsample the amount of information generated by the convolutional layer. This operation will half the size of the filters. Remember we start with 32 filters of 26x26, so after this operation will have 32 filters of 13x13.

We then take the (13, 13, 32) vector and flatten it to a (5408,) vector. Notice that 13 x 13

x 32 = 5408.

Finally, we add a couple more fully-connected layers (also called Dense layers.) Notice how the output layer has size 10 (one for each of our possible digit values) and a softmax activation. Softmax ensures we get a probability distribution indicating the most likely digit in the image.

```
model = Sequential([
    Conv2D(32, (3, 3), activation="relu",
input_shape=(28, 28, 1)),
    MaxPooling2D((2, 2)),
    Flatten(),
    Dense(100, activation="relu"),
    Dense(10, activation="softmax")
])
```

We now have our model. The next step is to define how we want to train it:

Let's use an SGD optimizer (Stochastic Gradient Descent) with 0.01 as the learning rate.

As we discussed before, we need to use the `sparse_categorical_crossentropy` loss because our target values are represented as integers. And we are going to compute the accuracy of our model as we train it. Notice in the summary of the model the shape of the vectors as they move through the layers we defined. They should look familiar after reading the explanation of our model above.

```
optimizer = SGD(learning_rate=0.01, momentum=0.9)
model.compile(
    optimizer=optimizer,
    loss="sparse_categorical_crossentropy",
    metrics=["accuracy"]
)
model.summary()

Model: "sequential_1"
```

Layer (type)	Output Shape	Param #
=====		

conv2d_1 (Conv2D)	(None, 26, 26, 32)	320
max_pooling2d_1 (MaxPooling2)	(None, 13, 13, 32)	0
flatten_1 (Flatten)	(None, 5408)	0
dense_2 (Dense)	(None, 100)	540900
dense_3 (Dense)	(None, 10)	1010
Total params: 542,230		
Trainable params: 542,230		
Non-trainable params: 0		

#Training and testing the model

At this point we are ready to fit our model on the train set.

For this example, we are going to run batches of 32 samples through our model for 10 iterations (epochs.) This should be enough to get a model with good predictive capabilities.

Note: This network is fairly shallow so it shouldn't take a long time to train it on a CPU. If you have access to a GPU it should be much faster.

```
model.fit(X_train, y_train, epochs=10, batch_size=32)
```

Epoch 1/10

1875/1875 [=====] - 7s 3ms/step - loss: 0.2412

- accuracy: 0.9252

Epoch 2/10

1875/1875 [=====] - 6s 3ms/step - loss: 0.0847 -

accuracy: 0.9748

Epoch 3/10

1875/1875 [=====] - 6s 3ms/step - loss: 0.0535 -

accuracy: 0.9831

Epoch 4/10

1875/1875 [=====] - 6s 3ms/step - loss: 0.0383 -

accuracy: 0.9886

Epoch 5/10

**1875/1875 [=====] - 6s 3ms/step - loss: 0.0288 -
accuracy: 0.9909**

Epoch 6/10

**1875/1875 [=====] - 6s 3ms/step - loss: 0.0212 -
accuracy: 0.9936**

Epoch 7/10

**1875/1875 [=====] - 6s 3ms/step - loss: 0.0163 -
accuracy: 0.9951**

Epoch 8/10

**1875/1875 [=====] - 6s 3ms/step - loss: 0.0129 -
accuracy: 0.9961**

Epoch 9/10

**1875/1875 [=====] - 6s 3ms/step - loss: 0.0092 -
accuracy: 0.9973**

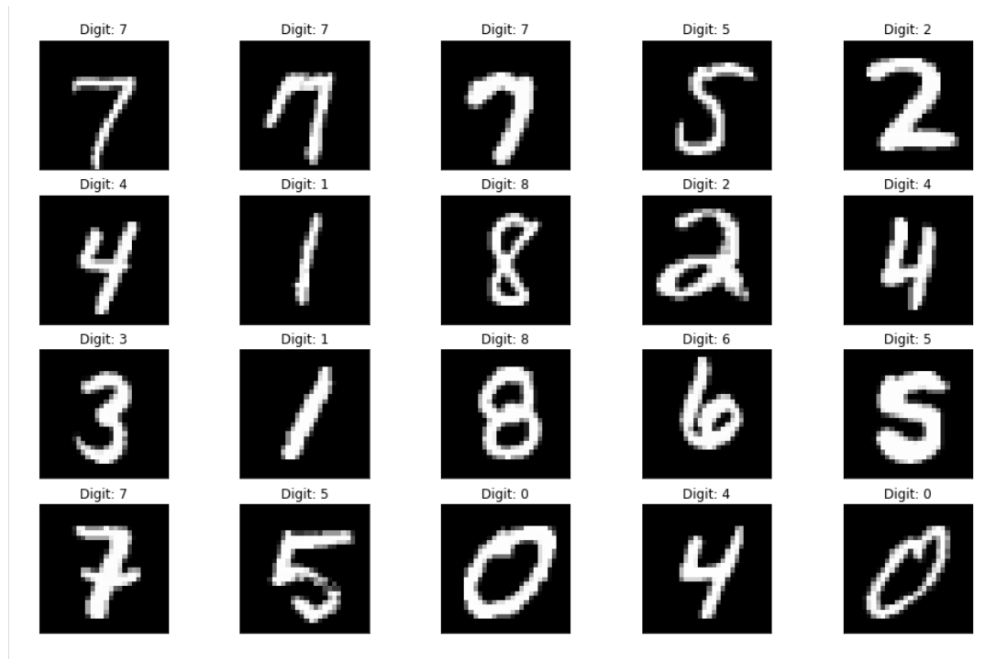
Epoch 10/10

**1875/1875 [=====] - 6s 3ms/step - loss: 0.0072 -
accuracy: 0.9983**

<keras.callbacks.History at 0x7ff9e24c1a20>

At this point you should have a model that scored above 99% accuracy on the train set. Now it's time to test it with a few of the images that we set aside on our test set. Let's run 20 random images through the model and display them together with the predicted digit:

```
plt.figure(figsize=(16, 10))
for i in range(20):
    image = random.choice(X_test).squeeze()
    digit = np.argmax(model.predict(image.reshape((1, 28,
28, 1))))[0], axis=-1)
    plot_digit(image, digit, plt, i)
plt.show()
```



The results look pretty good!

To get a much better idea about the quality of the predictions, we can run the entire test set (10,000 images) through the model and compute the final accuracy. To do this we can use the `accuracy_score()` function from SciKit-Learn passing a couple of arguments:

True values: The correct digit expected for each image. These are the values we have stored in `y_test`.

Predicted values: The predictions that our model made. These are the results of our model.

The final accuracy will be the value printed after running the cell.

```
predictions = np.argmax(model.predict(X_test), axis=-1)
accuracy_score(y_test, predictions)
0.9855
```

Estimating the model's performance

Now the trained model needs to be evaluated in terms of performance.

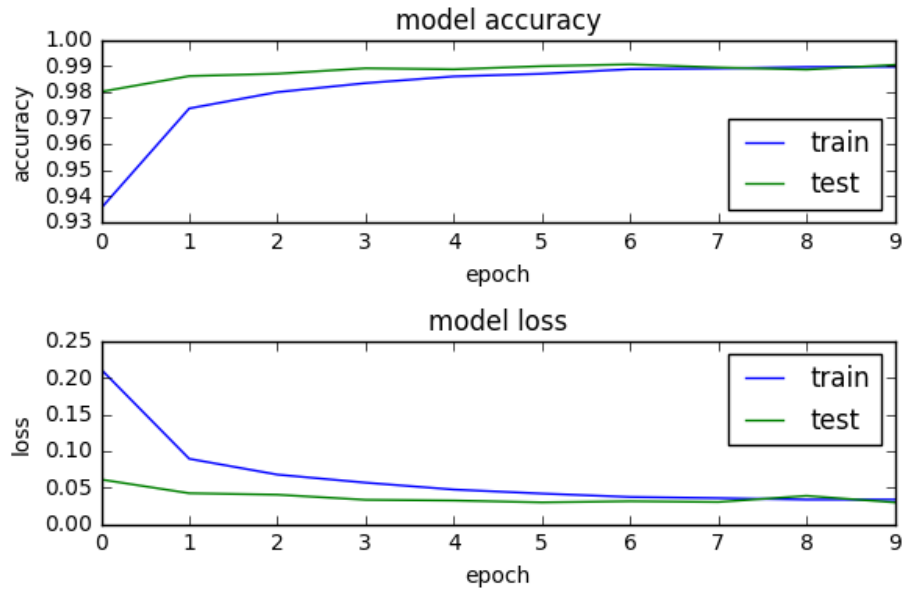
```
score = model.evaluate(X_test, y_test, verbose=0)
print('Test loss:', score[0]) #Test loss:
0.0296396646054
```

```
print('Test accuracy:', score[1]) #Test accuracy:
0.9904
```

Test accuracy 99%+ implies the model is trained well for prediction. If we visualize the whole training log, then with more epochs the loss and accuracy of the model on training and testing data converges thus making the model a stable one.

```
import os
# plotting the metrics
fig = plt.figure()
plt.subplot(2,1,1)
plt.plot(model_log.history['acc'])
plt.plot(model_log.history['val_acc'])
plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='lower
right')plt.subplot(2,1,2)
plt.plot(model_log.history['loss'])
plt.plot(model_log.history['val_loss'])
```

```
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upperright')
plt.tight_layout()fig
```



Saving the model to disk for reuse

Now, the trained model needs to be serialized. The architecture or structure of the model will be stored in a json file and the weights will be stored in hdf5 file format.

```
#Save the model
# serialize model to JSON
model_digit_json = model.to_json()
with open("model_digit.json", "w") as json_file:
    json_file.write(model_digit_json)
# serialize weights to HDF5
model.save_weights("model_digit.h5")
print("Saved model to disk")
```

Hence, the saved model can be reused later or easily ported to other environments too

Conclusion:

Thus, we have implemented the Image classification model using CNN. With the above code we can see that sufficient accuracy has been met. Throughout the epochs, our model accuracy increases and loss decreases that is good since our model gains confidence with our prediction

This indicates the model is trained in a good way

1. The loss is decreasing and the accuracy is increasing with every epoch.
2. The test accuracy is the measure of how good the model is predicting so, it is observed that the model is well trained after 10 epochs

References:

1. <https://www.analyticsvidhya.com/blog/2021/06/image-classification-using-convolutional-neural-network-with-python/>
2. Josh Patterson, Adam Gibson "Deep Learning: A Practitioner's Approach", O'Reilly Media, 2017
3. <https://deepnote.com/@svpino/MNIST-with-Convolutional-Neural-Network-s-a4b3c412-b802-4185-9806-02640fbba02e>
4. <https://towardsdatascience.com/a-simple-2d-cnn-for-mnist-digit-recognition-a998dbc1e79a>

Assignment No. 04

Problem Statement

Use Autoencoder to implement anomaly detection. Build the model by using

- a. Import required libraries
- b. Upload/access the dataset
- c. Encoder converts it into latent representation
- d. Decoder networks convert it back to the original input
- e. Compile the models with Optimizer, Loss, and Evaluation

Solution Expected

AutoEncoders are widely used in anomaly detection. The reconstruction errors are used as the anomaly scores. Let us look at how we can use AutoEncoder for anomaly detection using TensorFlow.

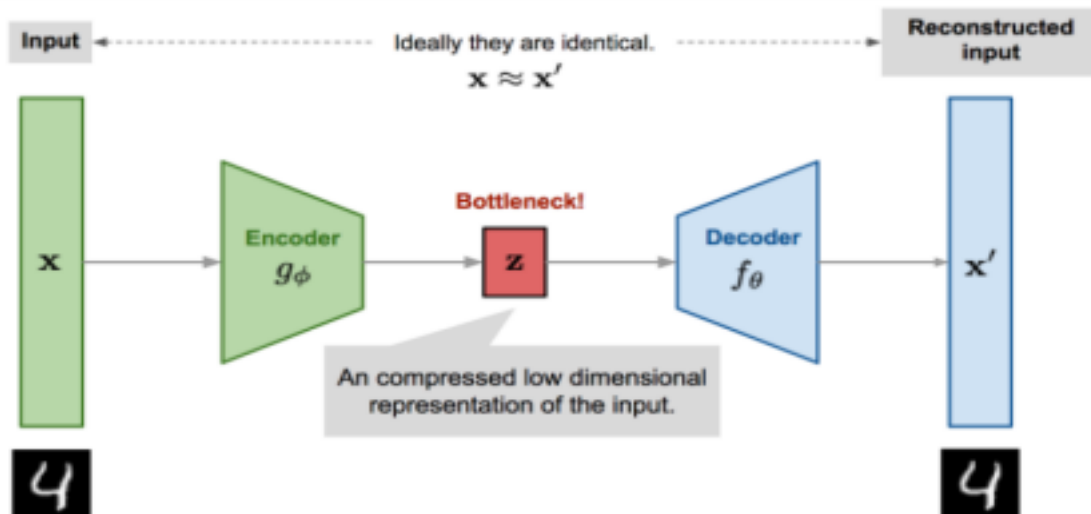
Import the required libraries and load the data. Here we are using the ECG data which consists of labels 0 and 1. Label 0 denotes the observation as an anomaly and label 1 denotes the observation as normal.

Objectives to be achieved

- 1) Use Autoencoder to implement anomaly detection.

Methodology to be used

AutoEncoder is a generative unsupervised deep learning algorithm used for reconstructing high-dimensional input data using a neural network with a narrow bottleneck layer in the middle which contains the latent representation of the input data.



Import required libraries

```
import pandas as pd
import numpy as np
import tensorflow as tf
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split

from sklearn.preprocessing import StandardScaler
from sklearn.metrics import confusion_matrix, recall_score,
accuracy_score, precision_score

RANDOM_SEED = 2021
TEST_PCT = 0.3
LABELS = ["Normal", "Fraud"]
```

Read the dataset

I had downloaded the data from Kaggle and stored it in the local directory.

```
dataset = pd.read_csv("creditcard.csv")
```

Exploratory Data Analysis

```
#check for any nullvalues
print("Any nulls in the dataset
",dataset.isnull().values.any() )
print('-----')
print("No. of unique labels ",
len(dataset['Class'].unique()))
print("Label values
",dataset.Class.unique())

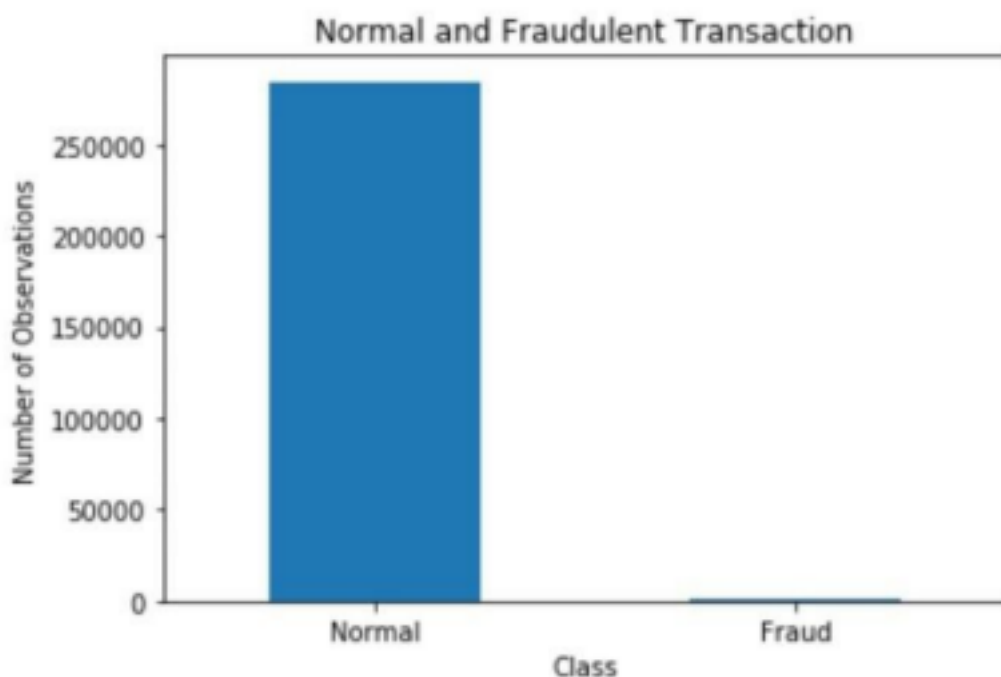
#0 is for normal credit card transaction
#1 is for fraudulent credit card
transaction
print('-----')
print("Break down of the Normal and Fraud
Transactions")
print(pd.value_counts(dataset['Class'],
sort = True) )
```

```
Any nulls in the dataset  False
-----
No. of unique labels  2
Label values  [0 1]
-----
Break down of the Normal and Fraud Transactions
0      284315
1         492
Name: Class, dtype: int64
```


Visualize the dataset

plotting the number of normal and fraud transactions in the dataset.

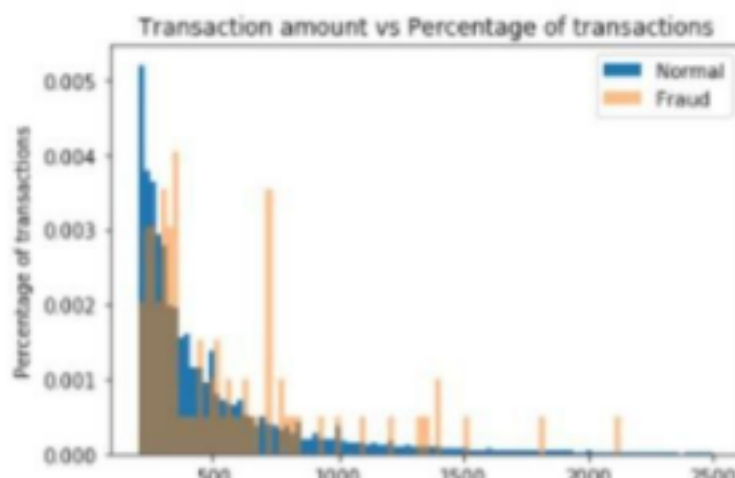
```
#Visualizing the imbalanced dataset
count_classes =
pd.value_counts(dataset['Class'], sort =
True)
count_classes.plot(kind = 'bar', rot=0)
plt.xticks(range(len(dataset['Class'].uni
que())), dataset.Class.unique())
plt.title("Frequency by observation
number")
plt.xlabel("Class")
plt.ylabel("Number of Observations");
```



Visualizing the amount for normal and fraud transactions.

```
# Save the normal and fraudulent transactions in separate dataframe
normal_dataset = dataset[dataset.Class == 0]
fraud_dataset = dataset[dataset.Class == 1]

#Visualize transaction amounts for normal and fraudulent transactions
bins = np.linspace(200, 2500, 100)
plt.hist(normal_dataset.Amount, bins=bins, alpha=1, density=True, label='Normal')
plt.hist(fraud_dataset.Amount, bins=bins, alpha=0.5, density=True, label='Fraud')
plt.legend(loc='upper right')
plt.title("Transaction amount vs Percentage of transactions")
plt.xlabel("Transaction amount (USD)")
plt.ylabel("Percentage of transactions");
plt.show()
```



Create train and test dataset

Checking on the dataset

V6	V7	V8	V9	V10	...	V22	V23	V24	V25	V26	V27	V28	Time	Amount	Class
0.482388	0.236999	0.098998	0.383787	0.090794	...	0.277838	-0.110474	0.089928	0.128839	-0.189115	0.133888	-0.021053	0.0	149.82	0
-0.082381	-0.078803	0.085102	-0.255425	-0.188974	...	-0.838872	0.101288	-0.338848	0.187170	0.125895	-0.008983	0.014724	0.0	2.89	0
1.800499	0.791481	0.247878	-1.514884	0.207843	...	0.771879	0.909412	-0.889281	-0.327842	-0.138097	-0.055353	-0.058752	1.0	378.88	0
1.247203	0.237899	0.377438	-1.387024	-0.084982	...	0.086274	-0.190321	-1.175575	0.847378	-0.221929	0.062723	0.081458	1.0	123.50	0
0.099921	0.892841	-0.279533	0.817739	0.783074	...	0.798278	-0.137488	0.141287	-0.208610	0.862292	0.219422	0.218153	2.0	89.99	0

Time and Amount are the columns that are not scaled, so applying StandardScaler to only Amount and Time columns. Normalizing the values between 0 and 1 did not work great for the dataset.

```
sc=StandardScaler()
dataset['Time'] =
sc.fit_transform(dataset['Time'].values.r
eshape(-1, 1))
dataset['Amount'] =
sc.fit_transform(dataset['Amount'].values
.reshape(-1, 1))
```

The last column in the dataset is our target variable.

The last column in the dataset is our target variable.

```
raw_data = dataset.values
# The last element contains if the
transaction is normal which is
represented by a 0 and if fraud then 1
labels = raw_data[:, -1]

# The other data points are the
electrocadiogram data
data = raw_data[:, 0:-1]

train_data, test_data, train_labels,
test_labels = train_test_split(
    data, labels, test_size=0.2,
    random_state=2021
)
```

Normalise the data to have a value between 0 and 1

Normalize the data to have a value between 0 and 1

```
min_val = tf.reduce_min(train_data)
max_val = tf.reduce_max(train_data)

train_data = (train_data - min_val) /
(max_val - min_val)
test_data = (test_data - min_val) /
(max_val - min_val)

train_data = tf.cast(train_data,
tf.float32)
test_data = tf.cast(test_data,
tf.float32)
```

Use only normal transactions to train the Autoencoder.

Normal data has a value of 0 in the target variable. Using the target variable to create a normal and fraud dataset.

```
train_labels = train_labels.astype(bool)
test_labels = test_labels.astype(bool)

#creating normal and fraud datasets
normal_train_data =
train_data[~train_labels]
normal_test_data =
test_data[~test_labels]

fraud_train_data =
train_data[train_labels]
fraud_test_data = test_data[test_labels]
print(" No. of records in Fraud Train
Data=",len(fraud_train_data))
print(" No. of records in Normal Train
data=",len(normal_train_data))
print(" No. of records in Fraud Test
Data=",len(fraud_test_data))
print(" No. of records in Normal Test
data=",len(normal_test_data))
```

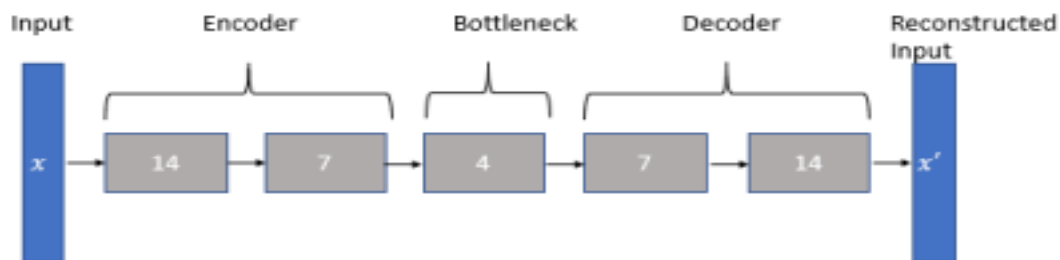
Set the training parameter values

Set the training parameter values

```
nb_epoch = 50
batch_size = 64
input_dim = normal_train_data.shape[1]
#num of columns, 30
encoding_dim = 14
hidden_dim_1 = int(encoding_dim / 2) #
hidden_dim_2=4
learning_rate = 1e-7
```

Create the Autoencoder

The architecture of the autoencoder is shown below.



```

#input Layer
input_layer =
tf.keras.layers.Input(shape=(input_dim,
))

#Encoder

encoder =
tf.keras.layers.Dense(encoding_dim,
activation="tanh",
activity_regularizer=tf.keras.regularizer
s.l2(learning_rate))(input_layer)
encoder=tf.keras.layers.Dropout(0.2)
(encoder)
encoder =
tf.keras.layers.Dense(hidden_dim_1,
activation='relu')(encoder)
encoder =
tf.keras.layers.Dense(hidden_dim_2,
activation=tf.nn.leaky_relu)(encoder)

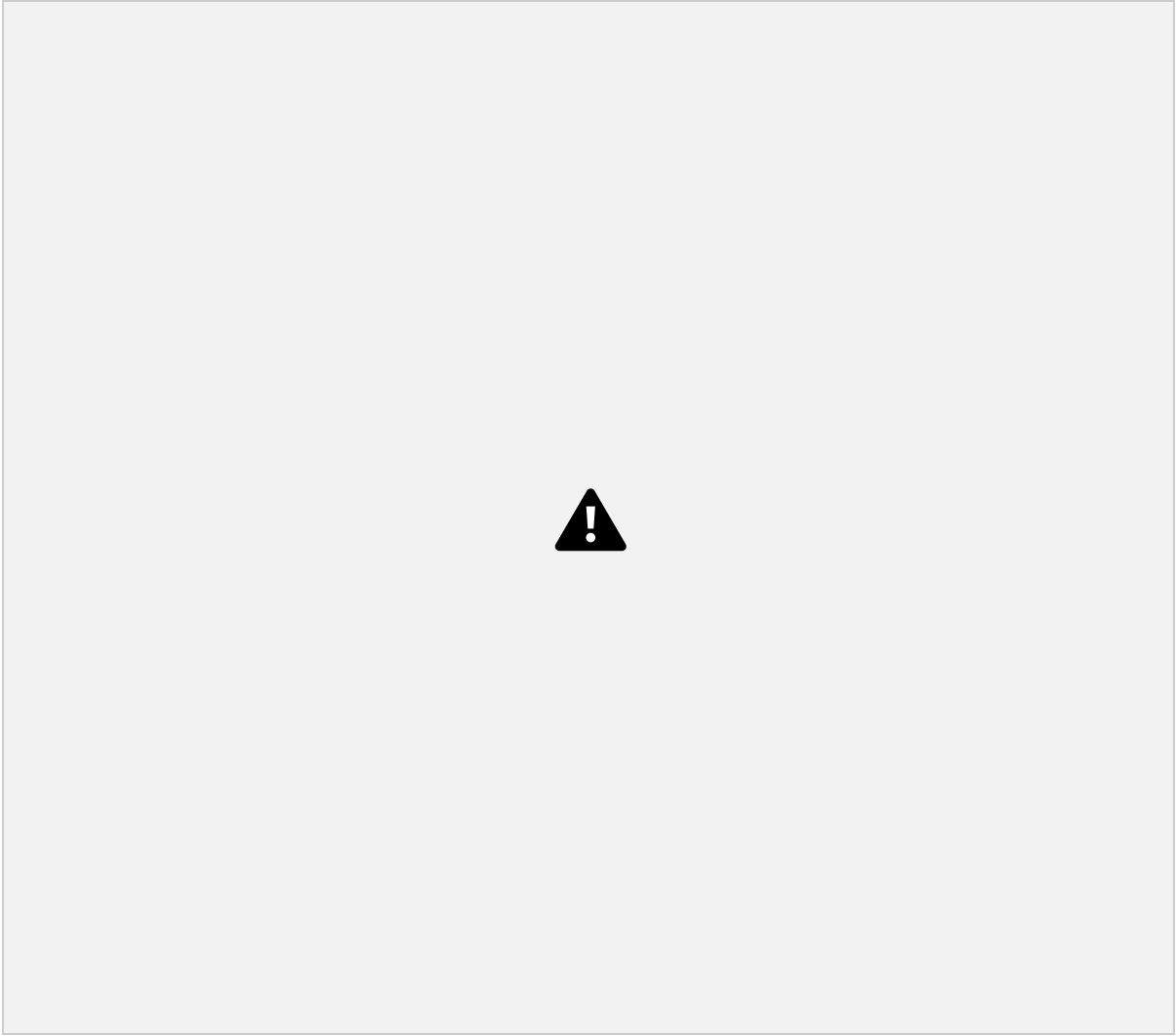
# Decoder
decoder =
tf.keras.layers.Dense(hidden_dim_1,
activation='relu')(encoder)
decoder=tf.keras.layers.Dropout(0.2)
(decoder)
decoder =
tf.keras.layers.Dense(encoding_dim,
activation='relu')(decoder)
decoder =
tf.keras.layers.Dense(input_dim,
activation='tanh')(decoder)

#Autoencoder
autoencoder =
tf.keras.Model(inputs=input_layer,
outputs=decoder)
autoencoder.summary()

```


Model: "functional_1"

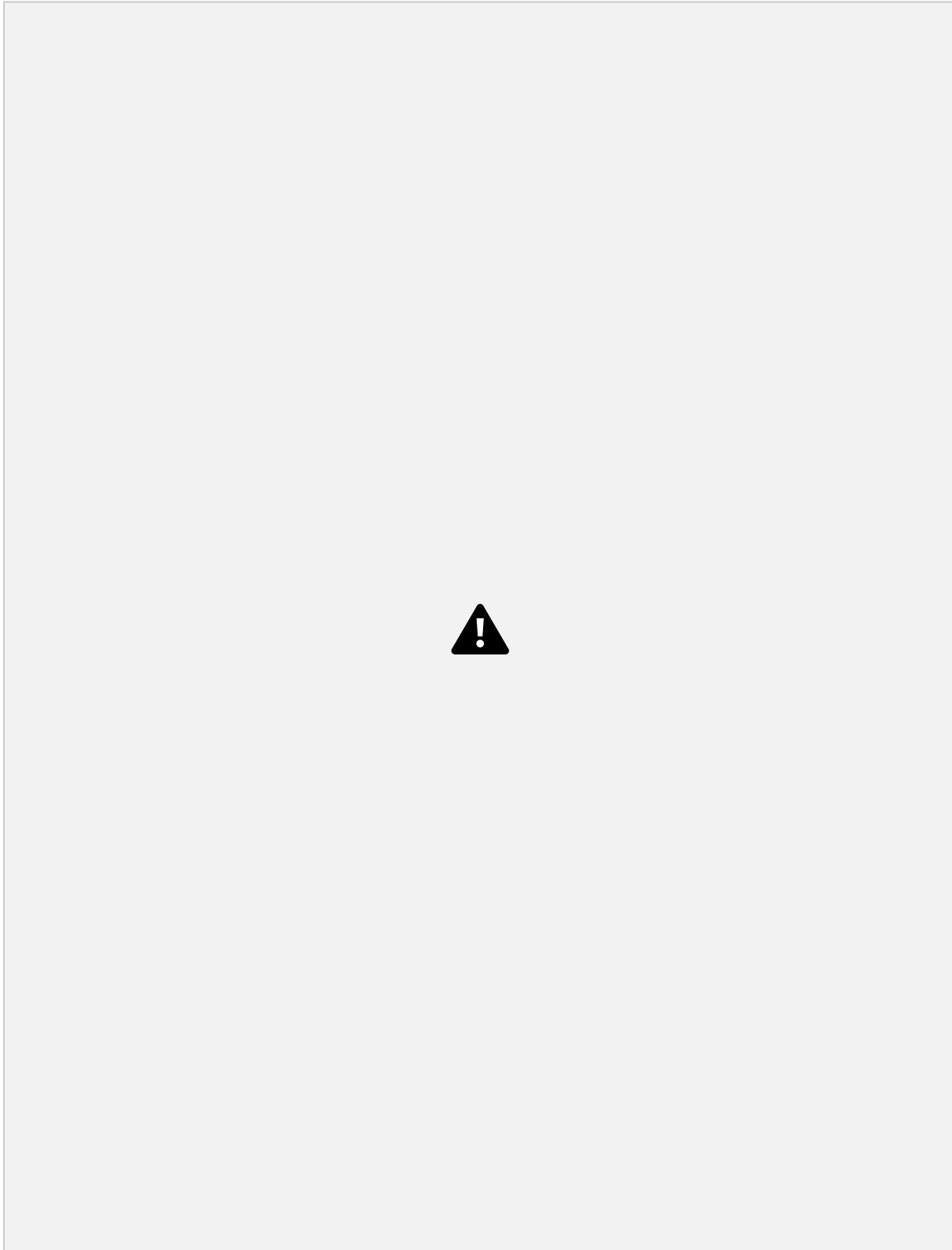
Layer (type)	Output Shape	Param #
=====		
input_1 (InputLayer)	[(None, 30)]	0
dense (Dense)	(None, 14)	434
dropout (Dropout)	(None, 14)	0
dense_1 (Dense)	(None, 7)	105
dense_2 (Dense)	(None, 4)	32
dense_3 (Dense)	(None, 7)	35
dropout_1 (Dropout)	(None, 7)	0
dense_4 (Dense)	(None, 14)	112
dense_5 (Dense)	(None, 30)	450
=====		
Total params: 1,168		
Trainable params: 1,168		
Non-trainable params: 0		



Plot training and test loss



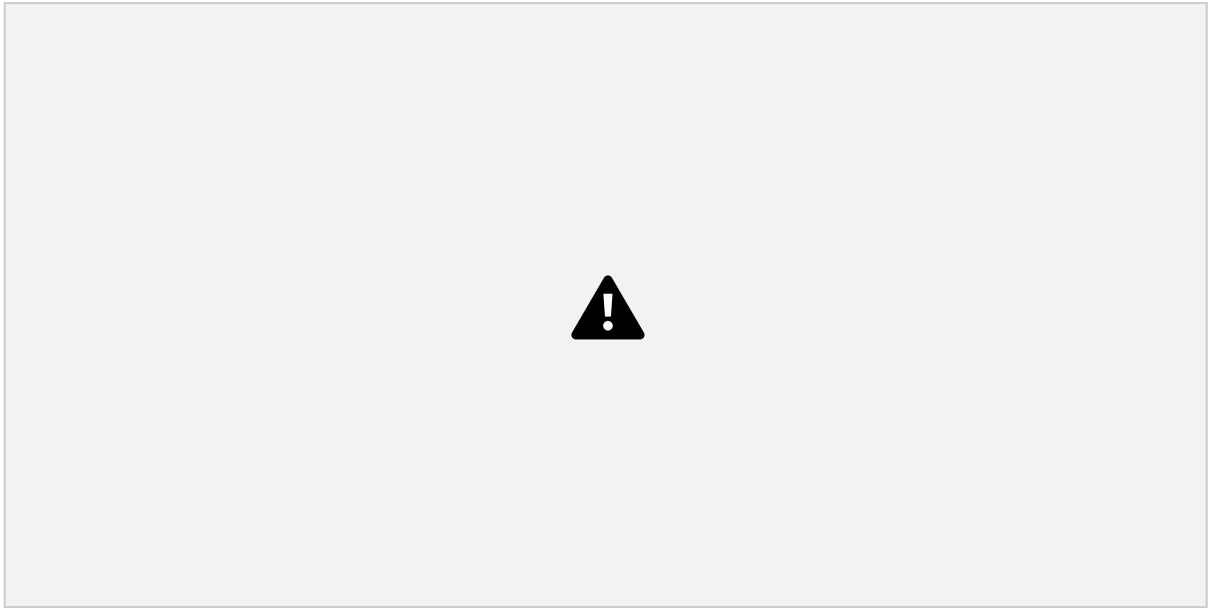
Plot training and test loss



Detect Anomalies on test data

Anomalies are data points where the reconstruction loss is higher

To calculate the reconstruction loss on test data, predict the test data and calculate the mean square error between the test data and the reconstructed test data.



Plotting the test data points and their respective reconstruction error sets a threshold value to visualize if the threshold value needs to be adjusted.



Detect anomalies as points where the reconstruction loss is greater than a fixed threshold. Here we see that a value of 52 for the threshold will be good.







As our dataset is highly imbalanced, we see a high accuracy but a low recall and precision.

Things to further improve precision and recall would add more relevant features, different architecture for autoencoder, different hyperparameters, or a different algorithm.

Conclusion:

Autoencoders can be used as an anomaly detection algorithm when we have an unbalanced dataset where we have a lot of good examples and only a few anomalies. Autoencoders are trained to minimize reconstruction error. When we train the autoencoders on normal data or good data, we can hypothesize that the anomalies will have higher reconstruction errors than the good or normal data.

The Continuous Bag Of Words (CBOW) Model in NLP Hands-On Implementation With Codes

Word2vec is considered one of the biggest breakthroughs in the development of natural language processing. The reason behind this is because it is easy to understand and use. Word2vec is basically a word embedding technique that is used to convert the words in the dataset to vectors so that the machine understands. Each unique word in your data is assigned to a vector and these vectors vary in dimensions depending on the length of the word.

The word2vec model has two different architectures to create the word embeddings. They are:

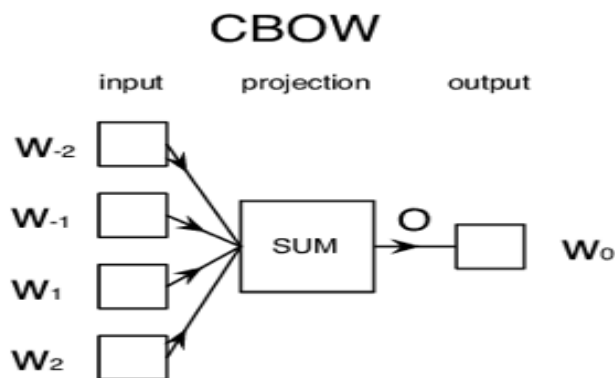
1. Continuous bag of words(CBOW)
2. Skip-gram model

In this article, we will learn about what CBOW is, the model architecture and the implementation of a CBOW model on a custom dataset.

What is the CBOW Model?

The CBOW model tries to understand the context of the words and takes this as input. It then tries to predict words that are contextually accurate. Let us consider an example for understanding this. Consider the sentence: 'It is a pleasant day' and the word 'pleasant' goes as input to the neural network. We are trying to predict the word 'day' here. We will use the one-hot encoding for the input words and measure the error rates with the one-hot encoded target word. Doing this will help us predict the output based on the word with least error.

Model Architecture



The CBOW model architecture is as shown above. The model tries to predict the target word by trying to understand the context of the surrounding words. Consider the same sentence as above, 'It is a pleasant day'. The model converts this sentence into word pairs in the form (contextword, targetword). The user will have to set the window size. If the window for the context word is 2 then the word pairs would look like this: ([it, a], is), ([is, pleasant], a), ([a, day], pleasant). With these word pairs, the model tries to predict the target word considered the context words.

If we have 4 context words used for predicting one target word the input layer will be in the form of four $1 \times W$ input vectors. These input vectors will be passed to the hidden layer where it is multiplied by a $W \times N$ matrix. Finally, the $1 \times N$ output from the hidden layer enters the sum layer where an element-wise summation is performed on the vectors before a final activation is performed and the output is obtained.

Implementation of the CBOW Model

For the implementation of this model, we will use a sample text data about coronavirus. You can use any text data of your choice.

The speed of transmission is an important point of difference between the two viruses. Influenza has a shorter median incubation period (the time from infection to appearance of symptoms) and a shorter serial interval (the time between successive cases) than COVID-19 virus. The serial interval for COVID-19 virus is estimated to be 5-6 days, while for influenza virus, the serial interval is 3 days. This means that influenza can spread faster than COVID-19.

Further, transmission in the first 3-5 days of illness, or potentially pre-symptomatic transmission - transmission of the virus before the appearance of symptoms - is a major driver of transmission for influenza. In contrast, while we are learning that there are people who can shed COVID-19 virus 24-48 hours prior to symptom onset, at present, this does not appear to be a major driver of transmission.

The reproductive number - the number of secondary infections generated from one infected individual - is understood to be between 2 and 2.5 for COVID-19 virus, higher than for influenza. However, estimates for both COVID-19 and influenza viruses are very context and time-specific, making direct comparisons more difficult.

```
import numpy as np
```

```
import keras.backend as K
```

```
from keras.models import Sequential
```

```
from keras.layers import Dense, Embedding, Lambda

from keras.utils import np_utils

from keras.preprocessing import sequence

from keras.preprocessing.text import Tokenizer

import gensim

data=open('/content/gdrive/My Drive/covid.txt','r')

corona_data = [text for text in data if text.count(' ') >= 2]

vectorize = Tokenizer()

vectorize.fit_on_texts(corona_data)

corona_data = vectorize.texts_to_sequences(corona_data)

total_vocab = sum(len(s) for s in corona_data)

word_count = len(vectorize.word_index) + 1

window_size = 2
```

In the above code, I have also used the built-in method to tokenize every word in the dataset and fit our data to the tokenizer. Once that is done, we need to calculate the total number of words

and the total number of sentences as well for further use. As mentioned in the model architecture, we need to assign the window size and I have assigned it to 2.

The next step is to write a function that generates pairs of the context words and the target words. The function below does exactly that. Here we have generated a function that takes in window sizes separately for target and the context and creates the pairs of contextual words and target words.

```
def cbow_model(data, window_size, total_vocab):

    total_length = window_size*2

    for text in data:

        text_len = len(text)

        for idx, word in enumerate(text):

            context_word = []

            target      = []

            begin = idx - window_size

            end = idx + window_size + 1

            context_word.append([text[i] for i in range(begin,
end) if 0 <= i < text_len and i != idx])
```

```
target.append(word)
```

```
contextual = sequence.pad_sequences(context_word,  
total_length=total_length)
```

```
final_target = np_utils.to_categorical(target, total_vocab)
```

```
yield(contextual, final_target)
```

Finally, it is time to build the neural network model that will train the CBOW on our sample data.

```
model = Sequential()
```

```
model.add(Embedding(input_dim=total_vocab, output_dim=100,  
input_length=window_size*2))
```

```
model.add(Lambda(lambda x: K.mean(x, axis=1),  
output_shape=(100,)))
```

```
model.add(Dense(total_vocab, activation='softmax'))
```

```
model.compile(loss='categorical_crossentropy',  
optimizer='adam')
```

```
for i in range(10):
```

```
cost = 0
```



```

for x, y in cbow_model(data, window_size, total_vocab):

    cost += model.train_on_batch(contextual, final_target)

print(i, cost)

```

Once we have completed the training its time to see how the model has performed and test it on some words. But to do this, we need to create a file that contains all the vectors. Later we can access these vectors using the gensim library.

```

dimensions=100

vect_file = open('/content/gdrive/My Drive/vectors.txt' , 'w')

vect_file.write('{} {} \n'.format(total_vocab,dimensions))

```

Next, we will access the weights of the trained model and write it to the above created file.

```

weights = model.get_weights()[0]

for text, i in vectorize.word_index.items():

    final_vec = ' '.join(map(str, list(weights[i, :])))

    vect_file.write('{} {} \n'.format(text, final_vec))

vect_file.close()

```

Now we will use the vectors that were created and use them in the gensim model. The word I have chosen is 'virus'.

```
cbow_output =  
gensim.models.KeyedVectors.load_word2vec_format('/content/gdrive/My  
Drive/vectors.txt', binary=False)
```

```
cbow_output.most_similar(positive=['virus'])
```

The output shows the words that are most similar to the word 'virus' along with the sequence or degree of similarity. The words like symptoms and incubation are contextually very accurate with the word virus which proves that CBOW model successfully understands the context of the data.

Conclusion:

In the above article, we saw what a CBOW model is and how it works. We also implemented the model on a custom dataset and got good output. The purpose here was to give you a high-level idea of what word embeddings are and how CBOW is useful. These can be used for text recognition, speech to text conversion etc.

References:

<https://analyticsindiamag.com/the-continuous-bag-of-words-cbow-model-in-nlp-hands-on-implementation-with-codes/>

Assignment No. 6

What Is Transfer Learning?

Transfer learning generally refers to a process where a model trained on one problem is used in some way on a second related problem.

In deep learning, transfer learning is a technique whereby a neural network model is first trained on a problem similar to the problem that is being solved. One or more layers from the trained model are then used in a new model trained on the problem of interest.

Transfer learning has the benefit of decreasing the training time for a neural network model and can result in lower generalization error.

The weights in re-used layers may be used as the starting point for the training process and adapted in response to the new problem. This usage treats transfer learning as a type of weight initialization scheme. This may be useful when the first related problem has a lot more labeled data than the problem of interest and the similarity in the structure of the problem may be useful in both contexts.

Transfer Learning in Keras with Computer Vision Models

Deep convolutional neural network models may take days or even weeks to train on very large datasets.

A way to short-cut this process is to re-use the model weights from pre-trained models that were developed for standard computer vision benchmark datasets, such as the ImageNet image recognition tasks. Top performing models can be downloaded and used directly, or integrated into a new model for your own computer vision problems.

- Transfer learning involves using models trained on one problem as a starting point on a related problem.
- Transfer learning is flexible, allowing the use of pre-trained models directly, as feature extraction preprocessing, and integrated into entirely new models.
- Keras provides convenient access to many top performing models on the ImageNet image recognition tasks such as VGG, Inception, and ResNet.



How to Use Transfer Learning when Developing Convolutional Neural Network Models

Overview

This tutorial is divided into five parts; they are:

1. What Is Transfer Learning?
2. Transfer Learning for Image Recognition
3. How to Use Pre-Trained Models
4. Models for Transfer Learning
5. Examples of Using Pre-Trained Models

What Is Transfer Learning?

Transfer learning generally refers to a process where a model trained on one problem is used in some way on a second related problem.

In deep learning, transfer learning is a technique whereby a neural network model is first trained on a problem similar to the problem that is being solved. One or more layers from the trained model are then used in a new model trained on the problem of interest.

This is typically understood in a supervised learning context, where the input is the same but the target may be of a different nature. For example, we may learn about one set of visual categories, such as cats and dogs, in the first setting, then learn about a different set of visual categories, such as ants and wasps, in the second setting.

Transfer learning has the benefit of decreasing the training time for a neural network model and can result in lower generalization error.

The weights in re-used layers may be used as the starting point for the training process and adapted in response to the new problem. This usage treats transfer learning as a type of weight initialization scheme. This may be useful when the first related problem has a lot more labeled data than the problem of interest and the similarity in the structure of the problem may be useful in both contexts.

... the objective is to take advantage of data from the first setting to extract information that may be useful when learning or even when directly making predictions in the second setting.

Transfer Learning for Image Recognition

A range of high-performing models have been developed for image classification and demonstrated on the annual [ImageNet Large Scale Visual Recognition Challenge](#), or ILSVRC.

This challenge, often referred to simply as [ImageNet](#), given the source of the image used in the competition, has resulted in a number of innovations in the architecture and training of convolutional neural networks. In addition, many of the models used in the competitions have been released under a permissive license. These models can be used as the basis for transfer learning in computer vision applications. This is desirable for a number of reasons, not least:

- **Useful Learned Features:** The models have learned how to detect generic features from photographs, given that they were trained on more than 1,000,000 images for 1,000

categories.

- **State-of-the-Art Performance:** The models achieved state of the art performance and remain effective on the specific image recognition task for which they were developed.
- **Easily Accessible:** The model weights are provided as free downloadable files and many libraries provide convenient APIs to download and use the models directly.

The model weights can be downloaded and used in the same model architecture using a range of different deep learning libraries, including Keras.

How to Use Pre-Trained Models

The use of a pre-trained model is limited only by your creativity.

For example, a model may be downloaded and used as-is, such as embedded into an application and used to classify new photographs.

Alternately, models may be downloaded and use as feature extraction models. Here, the output of the model from a layer prior to the output layer of the model is used as input to a new classifier model.

Recall that convolutional layers closer to the input layer of the model learn low-level features such as lines, that layers in the middle of the layer learn complex abstract features that combine the lower level features extracted from the input, and layers closer to the output interpret the extracted features in the context of a classification task.

Armed with this understanding, a level of detail for feature extraction from an existing pre-trained model can be chosen. For example, if a new task is quite different from classifying objects in photographs (e.g. different to ImageNet), then perhaps the output of the pre-trained model after the few layers would be appropriate. If a new task is quite similar to the task of classifying objects in photographs, then perhaps the output from layers much deeper in the model can be used, or even the output of the fully connected layer prior to the output layer can be used.

The pre-trained model can be used as a separate feature extraction program, in which case input can be pre-processed by the model or portion of the model to a given an output (e.g. vector of numbers) for each input image, that can then use as input when training a new model.

Alternately, the pre-trained model or desired portion of the model can be integrated directly into a new neural network model. In this usage, the weights of the pre-trained can be frozen so that they are not updated as the new model is trained. Alternately, the weights may be updated during the training of the new model, perhaps with a lower learning rate, allowing the pre-trained model to act like a weight initialization scheme when training the new model.

We can summarize some of these usage patterns as follows:

- **Classifier:** The pre-trained model is used directly to classify new images.
- **Standalone Feature Extractor:** The pre-trained model, or some portion of the model, is used to pre-process images and extract relevant features.
- **Integrated Feature Extractor:** The pre-trained model, or some portion of the model, is integrated into a new model, but layers of the pre-trained model are frozen during training.
- **Weight Initialization:** The pre-trained model, or some portion of the model, is integrated into a new model, and the layers of the pre-trained model are trained in concert with the new model.

Each approach can be effective and save significant time in developing and training a deep convolutional neural network model.

It may not be clear as to which usage of the pre-trained model may yield the best results on your new computer vision task, therefore some experimentation may be required.

Models for Transfer Learning

There are perhaps a dozen or more top-performing models for image recognition that can be downloaded and used as the basis for image recognition and related computer vision tasks.

Perhaps three of the more popular models are as follows:

- VGG (e.g. VGG16 or VGG19).
- GoogLeNet (e.g. InceptionV3).
- Residual Network (e.g. ResNet50).

These models are both widely used for transfer learning both because of their performance, but also

because they were examples that introduced specific architectural innovations, namely consistent and repeating structures (VGG), inception modules (GoogLeNet), and residual modules (ResNet).

Keras provides access to a number of top-performing pre-trained models that were developed for image recognition tasks.

They are available via the [Applications API](#), and include functions to load a model with or without the pre-trained weights, and prepare data in a way that a given model may expect (e.g. scaling of size and pixel values).

The first time a pre-trained model is loaded, Keras will download the required model weights, which may take some time given the speed of your internet connection. Weights are stored in the `.keras/models/` directory under your home directory and will be loaded from this location the next time that they are used.

When loading a given model, the `"include_top"` argument can be set to `False`, in which case the fully-connected output layers of the model used to make predictions is not loaded, allowing a new output layer to be added and trained. For example:

```
1 ...  
2 # load model without output layer  
3 model = VGG16(include_top=False)
```

Additionally, when the `"include_top"` argument is `False`, the `"input_tensor"` argument must be specified, allowing the expected fixed-sized input of the model to be changed. For example:

```
1 ...  
2 # load model and specify a new input shape for images  
3 new_input = Input(shape=(640, 480, 3))  
4 model = VGG16(include_top=False, input_tensor=new_input)
```

A model without a top will output activations from the last convolutional or pooling layer directly. One approach to summarizing these activations for their use in a classifier or as a feature vector representation of input is to add a global pooling layer, such as a max global pooling or average global pooling. The result is a vector that can be used as a feature descriptor for an input. Keras provides this capability directly via the

‘*pooling*’ argument that can be set to ‘*avg*’ or ‘*max*’. For example:

```
1 ...
2 # load model and specify a new input shape for images and avg pooling output
3 new_input = Input(shape=(640, 480, 3))
4 model = VGG16(include_top=False, input_tensor=new_input, pooling='avg')
```

Images can be prepared for a given model using the *preprocess_input()* function; e.g., pixel scaling is performed in a way that was performed to images in the training dataset when the model was developed. For example:

```
1 ...
2 # prepare an image
3 from keras.applications.vgg16 import preprocess_input
4 images = ...
5 prepared_images = preprocess_input(images)
```

Finally, you may wish to use a model architecture on your dataset, but not use the pre-trained weights, and instead initialize the model with random weights and train the model from scratch.

This can be achieved by setting the ‘*weights*’ argument to *None* instead of the default ‘*imagenet*’. Additionally, the ‘*classes*’ argument can be set to define the number of classes in your dataset, which will then be configured for you in the output layer of the model. For example:

```
1 ...
2 # define a new model with random weights and 10 classes
3 new_input = Input(shape=(640, 480, 3))
4 model = VGG16(weights=None, input_tensor=new_input, classes=10)
```

Now that we are familiar with the API, let’s take a look at loading three models using the Keras Applications API.

Load the VGG16 Pre-trained Model

The VGG16 model was developed by the Visual Graphics Group (VGG) at Oxford and was described in the 2014 paper titled “[Very Deep Convolutional Networks for Large-Scale Image Recognition.](#)”

By default, the model expects color input images to be rescaled to the size of 224×224 squares.

The model can be loaded as follows:

```
1 # example of loading the vgg16 model
2 from keras.applications.vgg16 import VGG16
3 # load model
4 model = VGG16()
5 # summarize the model
6 model.summary()
```

Running the example will load the VGG16 model and download the model weights if required.

The model can then be used directly to classify a photograph into one of 1,000 classes. In this case, the model architecture is summarized to confirm that it was loaded correctly.

References:

<https://machinelearningmastery.com/how-to-use-transfer-learning-when-developing-convolutional-neural-network-models/>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

<hr/>		
Layer (type)	Output Shape	Param #
=====		
input_1 (InputLayer)	(None, 224, 224, 3)	0
<hr/>		
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
<hr/>		
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
<hr/>		
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
<hr/>		
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
<hr/>		
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
<hr/>		
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
<hr/>		
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
<hr/>		
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
<hr/>		
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
<hr/>		
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
<hr/>		
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
<hr/>		

27			
28	block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
29			
30	block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
31			
32	block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
33			
34	block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
35			
36	block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
37			
38	block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
39			
40	block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
41			
42	flatten (Flatten)	(None, 25088)	0
43			
44	fc1 (Dense)	(None, 4096)	102764544
45			
46	fc2 (Dense)	(None, 4096)	16781312
47			
48	predictions (Dense)	(None, 1000)	4097000
49	=====		
50	Total params: 138,357,544		
51	Trainable params: 138,357,544		
52	Non-trainable params: 0		
53			

Load the InceptionV3 Pre-Trained Model

The InceptionV3 is the third iteration of the inception architecture, first developed for the GoogLeNet model.

This model was developed by researchers at Google and described in the 2015 paper titled “[Rethinking the Inception Architecture for Computer Vision](#).”

The model expects color images to have the square shape 299×299.

The model can be loaded as follows:

```
1 # example of loading the inception v3 model
2 from keras.applications.inception_v3 import InceptionV3
3 # load model
4 model = InceptionV3()
5 # summarize the model
6 model.summary()
```

Running the example will load the model, downloading the weights if required, and then summarize the model architecture to confirm it was loaded correctly.

The output is omitted in this case for brevity, as it is a deep model with many layers.

Load the ResNet50 Pre-trained Model

The Residual Network, or ResNet for short, is a model that makes use of the residual module involving shortcut connections.

It was developed by researchers at Microsoft and described in the 2015 paper titled “[Deep Residual Learning for Image Recognition](#).”

The model expects color images to have the square shape 224×224.

```
1 # example of loading the resnet50 model
2 from keras.applications.resnet50 import ResNet50
3 # load model
4 model = ResNet50()
5 # summarize the model
6 model.summary()
```

Running the example will load the model, downloading the weights if required, and then summarize the model architecture to confirm it was loaded correctly.

The output is omitted in this case for brevity, as it is a deep model