A Mini Project with Seminar On

## OBJECT RECOGNITION IN PHOTOGRAPHS USING CNN

Submitted in partial fulfillment of the requirements for the award of the

## Bachelor of Technology

In

## Department of Computer Science and Engineering

By

|  |  |
| --- | --- |
| **AKURATHI ARAVINDA** | **18241A0561** |
| **CHALLAGULLA YOSHITHA** | **18241A0569** |
| **KAKARLA MEGHANA** | **18241A0580** |

Under the Esteemed guidance of

**Dr. K.ANURADHA,**

**Professor and Dean of ICT**



**Department of Computer Science and Engineering**

**GOKARAJU RANGARAJU INSTITUTE OF ENGINEERING AND TECHNOLOGY (Approved by AICTE, Autonomous under JNTUH), Hyderabad, Bachupally, Kukatpally, Hyderabad-500090**



**GOKARAJU RANGARAJU INSTITUTE OF ENGINEERING AND TECHNOLOGY (Approved by AICTE, Autonomous under JNTUH), Hyderabad, Bachupally, Kukatpally, Hyderabad-500090**

**CERTIFICATE**

This is to certify that the mini project entitled “**Object Recognition in photographs using CNN**” is submitted by **A. Aravinda (18241A0561), CH. Yoshitha (18241A0569), K. Meghana (18241A0580)** in partial fulfillment of the award of degree in BACHELOR OF TECHNOLOGY in Computer Science and Engineering during academic year 2020-2021.

INTERNAL GUIDE HEAD OF THE DEPARTMENT

**Dr.K.ANURADHA Dr. K. MADHAVI**

Professor

EXTERNAL EXAMINER

## ACKNOWLEDGEMENT

There are many people who helped us directly and indirectly to complete our project successfully. We would like to take this opportunity to thank one and all. First we would like to express our deep gratitude towards our internal guide **Dr. K.Anuradha, Prof.** Department of CSE for her support in the completion of our project. We wish to express our sincere thanks to **Dr. K. Madhavi, HOD, Department of CSE** and to our principal **Dr. J. Praveen** for providing the facilities to complete the dissertation. We would like to thank all our faculty and friends for their help and constructive criticism during the project period. Finally, we are very much indebted to our parents for their moral support and encouragement to achieve goals.

**A.ARAVINDA (18241A0561)**

**CH. YOSHITHA (18241A0569)**

**K. MEGHANA (18241A0580)**

## DECLARATION

We hereby declare that the mini project entitled **“Object Recognition in Photographs using CNN”** is the work done during the period from **8th March 2021 to 8th July 2021** and is submitted in the partial fulfillment of the requirements for the award of degree of Bachelor of Technology in Computer Science and Engineering from Gokaraju Rangaraju Institute of Engineering and Technology (Autonomous under Jawaharlal Nehru Technology University, Hyderabad).The results embodied in this project have not been submitted to any other university or Institution for the award of any degree or diploma.

A.ARAVINDA (18241A0561)

CH. YOSHITHA (18241A0569)

K. MEGHANA (18241A0580)

## ABSTRACT

Artificial Intelligence (AI) has gone a long way, smoothly bridging the gap between human and machine potential. Data enthusiasts from all over the world are working on various aspects of AI to make their visions a reality, one of which is Deep Learning. This field aims to make it possible for robots to see the world in the same way that humans do, and to use that information for a variety of tasks and operations (such as image recognition, image analysis, and classification)Developments in Computer Vision using Deep Learning have been a significant success, notably with the Convolutional Neural Network approach.

Have you ever wondered how social media facial recognition works, how object detection aids in the development of self-driving cars, or how illness diagnosis in healthcare is done via visual imagery? Convolutional neural networks make it all possible (CNN). A convolutional neural network is a feed-forward neural network that is commonly used to analyse visual images. It's also referred to as a ConvNet..

The following is how CNN detects an image: The image's pixels are supplied into the convolutional layer, which conducts the convolution. The end result is a jumbled map. The convolved map is applied to a ReLU function to obtain a corrected feature map.. Multiple convolutions and ReLU layers are used to locate the features in the image. To identify specific parts of the image, various pooling layers with various filters are used.

To get the final output, the pooled feature map is flattened and fed to a fully connected layer. Image sharpening and restoration, which includes zooming, blurring, sharpening, grey scale to colour conversion, and detecting edges and vice versa, are some of the applications of Object recognition. One of the most common tasks is hurdle detection, which is accomplished by identifying various types of objects in an image and then calculating the distance between the robot and the hurdles.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| CONTENTS | | Page |
| No. |
| Title Page | | 1 |
| Declaration | | 4 |
| Certificate by the Supervisor | | 2 |
| Acknowledgement | | 3 |
| Abstract | | 5 |
| Chapter 1: Introduction .................................................................................................... | | 8 |
| 1.1 Rationale.................................................................................................................... | | 10 |
| 1.2 | Goal....................................................................................................................... | 11 |
| 1.3 | Objective................................................................................................................. | 11 |
| 1.4 | Methodology........................................................................................................... | 12 |
| 1.5 | Roles and Responsibilities...................................................................................... | 17 |
| 1.6 | Contribution of Project........................................................................................ | 18 |
| 1.6.1 Market Potential............................................................................................... | | 18 |
| 1.6.2 Innovativeness.................................................................................................. | | 18 |
| 1.6.3 Usefulness........................................................................................................ | | 19 |
| 1.7 | Report Organization.............................................................................................. | 20 |
| Chapter 2: Requirement Engineering................................................................................ | | 21 |
| 2.1 | Functional Requirement........................................................................................ | 21 |
| 2.1.1 Input and Output...................................................................................... | | 21 |
| 2.2 | Non Functional Requirements.............................................................................. | 21 |
| Chapter 3: Analysis & Design........................................................................................... | | 22 |
| 3.1 | Use-case Diagrams............................................................................................... | 22 |
| 3.2 | Activity Diagrams................................................................................................. | 23 |

|  |  |  |  |
| --- | --- | --- | --- |
| 3.3 | Class Diagram…………………………………………………………………... | | 24 |
| 3.4 | Sequence Diagram……………………………………………………………… | | 25 |
| 3.5 | System Architecture............................................................................................. | | 26 |
| Chapter 4: Construction.................................................................................................... | | | 27 |
| 4.1 | Implementation..................................................................................................... | | 27 |
| 4.2 | Implementation Details.......................................................................................... | | 35 |
| 4.3 | Software Details.................................................................................................... | | 39 |
| 4.4 | Hardware Details................................................................................................... | | 39 |
| 4.5 | Testing................................................................................................................... | | 40 |
| 4.5.1 | | White Box Testing........................................................................................... | 40 |
| 4.5.2 | | Black Box Testing ....................................................................................... | 41 |
| Chapter 5: Conclusion and Future scope....................................................................... | | | 44 |
| 5.1 Conclusion........................................................................................................... | | | 44 |
| 5.2 Future Scope........................................................................................................ | | | 45 |
| 5.3 Expected Outcome.............................................................................................. | | | 46 |
| References………………………………………..…………………………………..… | | | 47 |

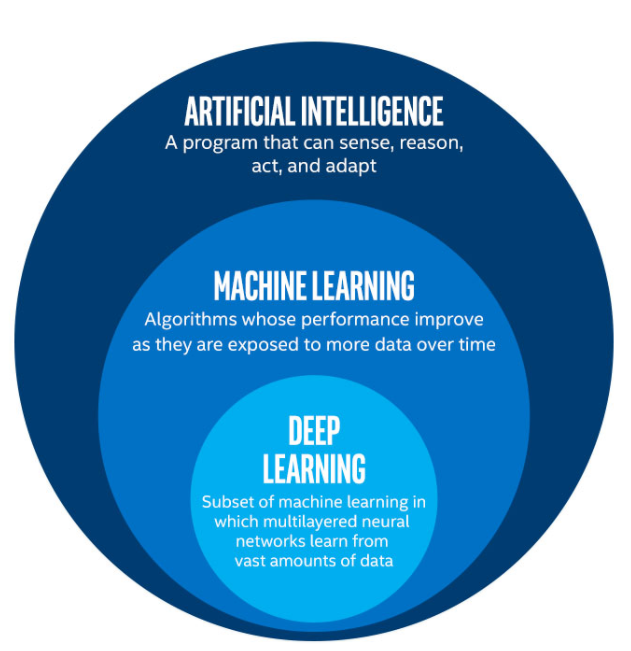
**Chapter-1**

**INTRODUCTION**

Artificial intelligence (AI) is a branch of computer science that focuses on the creation of intelligent systems or machines that function and react in the same way that humans do.Humans have a high ability to identify items through their vision, making it quite easy to discover and recognise them.

Deep Learning is the next step in the evolution of Machine Learning, and it is inspired by how the human brain works as shown in fig1.1. Deep Learning is used to solve complex problems involving large amounts of data that are diverse and unstructured. Artificial Neural Networks (ANNs) are used to build deep learning models, which mimic how the human brain works.

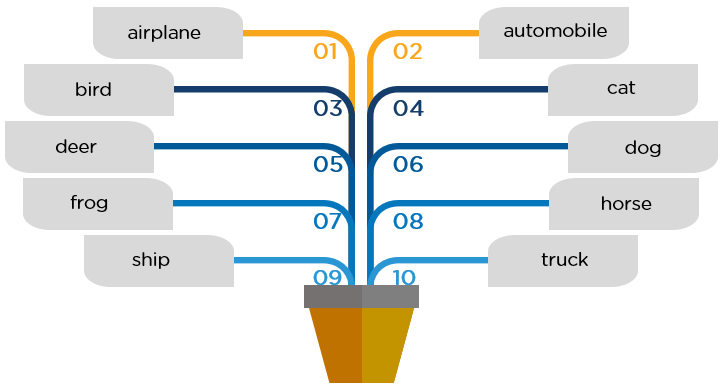
While computer-assisted deductive reasoning, inference, and decision-making are still a long way off, significant progress has been made in the application of AI techniques and algorithms The concepts presented here are complex and technical, and they are based on mathematics, statistics, probability theory, physics, signal processing, machine learning, computer science, and psychology.Object detection and recognition, on the other hand, is a major problem for machines. In the field of computer science, 'Neural networks' have been created to solve this problem. Artificial Neural Networks is another name for it. Non-symbolic artificial intelligence is represented by neural networks.

**Fig1.1** 

They are computer models of the human brain that aid in the detection and recognition of objects. Object detection and recognition, on the other hand, is a topic of study in which a lot of research is being done.

However, the focus of the study is on dynamic things, or moving objects. The system is designed to cope with static item detection and recognition. Under the probabilistic modelling of natural images, 32x32 pixel colour images from the CIFAR-10 dataset count.

The inputs to the algorithm are 32x32 pixels colour image of CIFAR-10 dataset which consists of the classes as shown in the below figure 1.2:



**Fig 1.2**

This dataset consists of 60,000 photos where each object is a photograph, out of which 50,000 are for training and 10,000 are for testing. And we classify these images using CNN.

**Rationale:**

Machine studying and sample recognition, each of which use convolutional neural networks, have made significant progress in recent years (CNNs). Image classification and object detection, both of which are part of the object identification field, are exciting challenges in computer vision.

Advances in complex neural networks, deep learning techniques and the increased parallel processing power provided by graphics processing units have contributed to the powerful scientific development of this field (GPU) in recent years. The set of fixed categories of input images is called the image classification problem.

Object recognition is advancing in a variety of industries, from personal security to office productivity. They continue to pose serious obstacles. The possibilities are endless when it comes to future object detection use cases Widely used to import data from printed data records such as passport documents, invoices, bank statements, computer receipts, business cards, letters, static data prints, and other documents. It can be used in machine processes such as digitizing and editing printed documents, searching, storing them electronically compactly, displaying them online, and receiving texts (extractions) from calculations, knowledge, machine translation, and voice. This is a common way to do this.

Image classification and object detection, which are part of the field of object recognition, are interesting issues in computer vision. In recent years it has contributed to a solid scientific development in these areas (GPU). The task of assigning labels to input images from a fixed set of categories is called an image classification problem. Includes tagging skin cancer images and the use of high-resolution images to detect natural disasters such as floods, volcanoes, and severe droughts and document their effects and damage.

The function used to power the image classification algorithm is critical to its performance. In summary, advances in machine learning-based image classification technology have relied heavily on techniques to select essential image functions from databases. These resources can be a daunting task, adding computational complexity and cost. Image classification, feature extraction, and selection of learning algorithms typically require two independent steps, which were largely developed and improved on in our project.

**Goal:**

We focus mainly on:

* Using convolutional neural networks to train the model for feature extraction and classification.
* Recognising the images in CIFAR-10 dataset.
* To recognize instances of a predefined set of object classes.
* Using the hidden layers, we used CNN for object detection to recognise patterns such as edges (vertical/horizontal), shapes, colours, and textures.

## Objective:

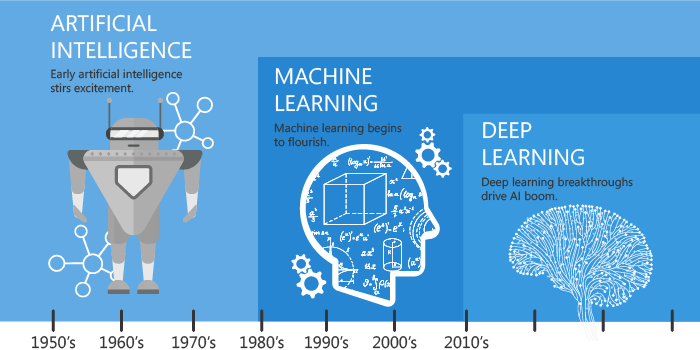
Our project's objectives are as follows

* Loading a data set
* To extract features using CNN hidden layers.
* Classification.
* Analysis.
* To compile and evaluate the model.
* To figure out what class an image belongs to.

**Methodology**

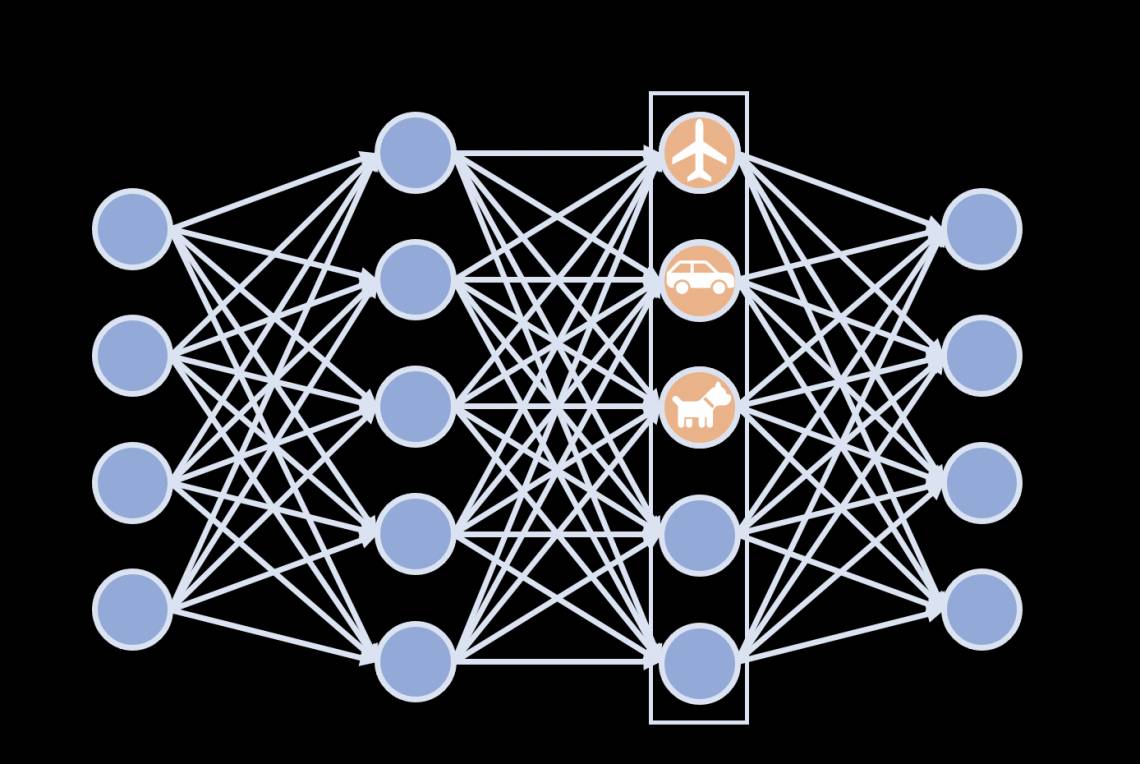
**Deep Learning**

Deep learning is a subset of machine learning in artificial intelligence that has networks capable of learning unsupervised from data that is unstructured or unlabelled.Deep neural network or deep neural learning are some other terms for about the same thing. Deep learning is a type of artificial intelligence that mimics the human brain's data processing to identify objects, recognise speech, translate languages, and make decisions. Smaller subsets of artificial intelligence-first machine learning Smaller subsets of artificial intelligence-first machine learning, then deep learning, a subset of machine learning-have created ever larger disruptions since an early flush of optimism in the 1950s. (as shown in the fig1.3).



**Fig1.3**

Assume you're a small child who comes across a pot of boiling water. You won't know it's hot unless someone tells you it is until you see it. When you touch a hot pot with your fingers, your nerves send the information to your brain, where neurons process it and send the signal back to your fingers, causing you to feel the heat. If you see a hot pot again, your brain will recall the previous incidents and remind you that touching it will make you feel hot. And, in order to make the best decision possible in any situation, our brain is constantly learning from the environment and previous experiences. This is essentially what Deep Learning accomplishes! It corrects itself without explicit programming as It gains knowledge from both raw data and prior experiences.. The below fig1.4 shows the neural network in deep learning.



**Fig1.4**

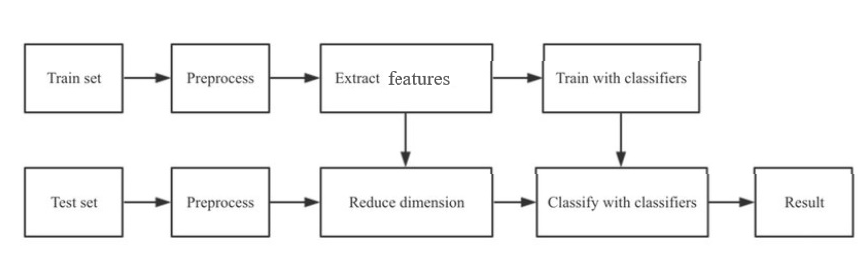
Learning at a deeper level AI can learn from both labelled and unlabeled data without the need for human intervention. Deep learning, a type of machine learning, could help with fraud and money laundering detection, among other things.The deep learning process contains the steps as shown in fig 1.5.

Deep learning has been successfully applied to a wide range of applications, and it is now considered one of the most cutting-edge machine learning and AI techniques. The associated algorithms are frequently used for supervised, unsupervised, and semi-supervised learning problems.

When machines can perform tasks that would normally require human intelligence, this is referred to as artificial intelligence. Machine learning, for example, is when machines learn through experience and acquire skills without the need for human intervention.

Similar to how we learn from experience, the deep learning algorithm would repeat a task, tweaking it slightly each time to improve the outcome.

We call it "deep learning" because neural networks have multiple (deep) layers that enable learning. Any problem that requires "thought" to solve can be learned to solve by deep learning.

Deep learning allows machines to solve complex problems despite being given a large, unstructured, and interconnected data set. Deep learning algorithms get better as they learn more.

**Fig 1.5-Deep learning process**

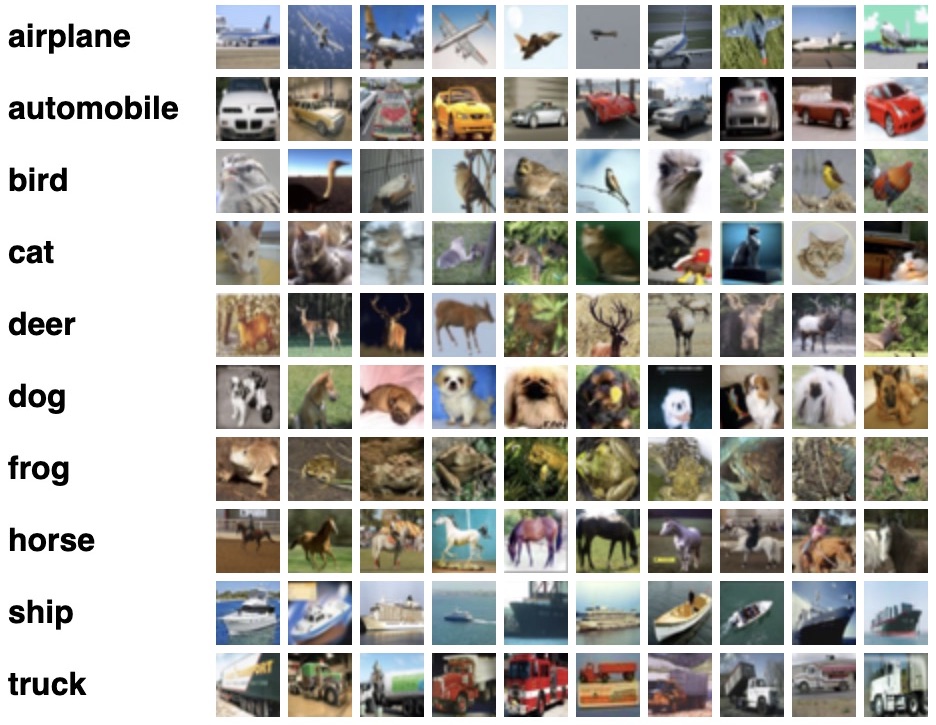
The inputs to our algorithms are 32x32 pixel color images. The classes and the standard integer values that go with them are listed below:

* 0- Airplane
* 1- Automobile
* 2- Bird
* 3- Cat
* 4- Deer
* 5- Dog
* 6- Frog
* 7- Horse
* 8- Ship
* 9- Truck

The recognition of the image's class is the output. We experimented with a variety of layers, including convolutional layers with various filters, pooling, flattening, dense layers with various activation functions, such as ReLu and SoftMax, and dense layers with various activation functions, such as ReLu and SoftMax, and dense layers with various activation functions, such as ReLu and SoftMax.

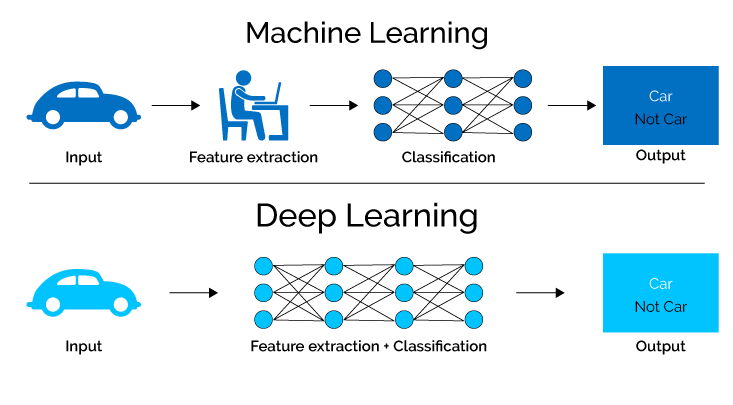
**CIFAR-10 dataset:**

Alex Krizhevsky, Vinod Nair and Geoffrey Hinton collected data sets from the Canadian Institute for Advanced Research (ICRA10). The CIFAR10 dataset contains 60,000 32x32 color images divided into 10 layers, each containing 6,000 images. Each of the 10,000 images in the dataset is divided into 5 training lots and 1 test lot. The test set consisted of exactly 1000 photos from each class, randomly selected. The rest of the photos were divided into learning lots in a random order. However, in some batches, one layer may contain more images than others. The training package consists of exactly 5000 photos of each layer between them. Figure 1.6 shows some images contained in the CIFAR10 dataset.



**Fig1.6**

**Methodology:**

The whole numerical values of the picture are not given into the CNN inputs. Instead, the whole image is divided into many smaller sets, each acting as its own image. The image is divided into small portions by a modest filter size. A little piece of the picture is linked to each group of neurons. These photos are then processed in the same way as a standard neural network. The computer collects patterns on the image and stores the results as a matrix. This operation is repeated until the system receives the complete image bit by bit. The outcome is a huge Matrix that represents the many patterns recorded by the system from the input image. These images are then processed in the same manner as a traditional neural network would. The computer collects the patterns associated with the image and stores them as a matrix. This process continues until the system receives the full-size bit image. The result is a large matrix representing a large number of system models extracted from the input image, as shown in Figure 1.7 below. Difference between machine learning and deep learning.

**Fig1.7**

**Role and Responsibilities:**

|  |  |
| --- | --- |
| **NAME** | **ROLE AND RESPONSIBILITIES** |
| A.Aravinda | * Data analyst * Machine Learning * Coding * Documentation |
| CH.Yoshitha | * Data analyst * Machine Learning * Coding * Documentation |
| K.Meghana | * Data analyst * Machine Learning * Implementation * Documentation |

**Contribution of Project**

* + 1. **Market potential**

Object recognition software is changing the way businesses engage with customers, market products online and offline, and manage store inventories. Object recognition technology is increasingly being used to find things for purchase on the internet.

For example, Snap, Snapchat's parent company, is working on a visual product feature in Snapchat that would employ object recognition technology to allow users to capture images of objects in the real world, as well as identify, explore, and purchase items on Amazon.

Entrepreneurs and merchants have understood that traditional sales promotion, marketing, and visual merchandising techniques will not be sustainable in the long run in the e-commerce industry, especially during the COVID-19 epidemic. As a result, merchants are rapidly adjusting to the new era of AI and object recognition in order to provide superior customer service.

The application of object recognition for shelf recognition, product placement, and merchandising standards compliance is gradually gaining momentum. Manufacturers, retailers, and marketers may use object recognition technology to better understand their markets and respond quickly.

Major cloud providers such as Amazon Web Services, Microsoft, and Google are spending extensively in improving their object recognition market capabilities in order to improve in-store and online retail execution. Due to the outbreak of COVID-19, demand for cloud-based services has increased even more.

## Innovativeness

Modeling a rich distribution with thousands or millions of random variables is a computationally and statistically difficult undertaking.

Let's say we merely want to model binary variables; this is the simplest situation, yet it's already overwhelming. There are 23072 potential binary pictures of this kind for a modest 32x32 pixels color (RGB) image. This number is about 400,410,800,400 times higher than the estimated number of atoms in the universe.

**Usefulness**

Using Object Recognition

* Drones: Object recognition drones can provide automated remote monitoring, inspection and control of assets based on vision.
* Production: Inspect equipment regularly. Analysis of production lines and key positions inside. Pay attention to the quality of the final item to reduce defects. Worker health assessment gives manufacturers full control over various system operations
* Self-driving cars: By recognizing objects in self-driving cars, you can detect activity on the road and take appropriate action. Mini robots help logistics to find objects and move them from one location to another. It also tracks product movement.
* Military surveillance: Anomalous activity detection and automated decision-making in border areas help prevent incursions and save soldiers' lives.
* Forestry Operations: Unmanned aerial vehicles (UAVs) can monitor forests, predict changes that could cause wildfires, and prevent poaching.It also has the potential to enable comprehensive monitoring of large areas that are not easily accessible to humans.



**Fig1.8**

Report Organization

The remaining section of the report is structured as follows:

* **Chapter 2** provides detailed business and technical requirements
* **Chapter 3** provides analysis and design of this project
* **Chapter 4** provides Construction, implementation details of this project
* **Chapter 5** provides Conclusion and future scope as well as future application of this project

# Chapter-2

# REQUIREMENT ENGINEERING

* 1. **Functional Requirement**

The functional requirements describe the basic functionality of the application.

**Input:** send index of range 0-9999 to input an image from CIFAR-10 dataset for predict function.

**Output:** if the index is within the range then, clicking enter gives the predicted output class label along with the input image. Else it returns the respective error.

## Non Functional Requirement

Non-functional needs are those that aren't directly related to the system's unique functionality. They might be linked to emergent features like dependability, usability, and so on.

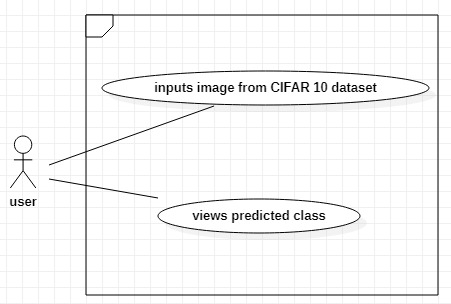
* Availability.
* Reliability.
* Maintainability.
* Easy to use.
* Provides visual analysis.

# Chapter-3

## ANALYSIS AND DESIGN

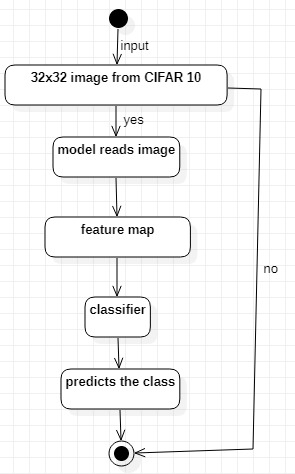
**Use case diagram**

Use case diagrams are a type of diagram that is used to collect information about a system's requirements, covering both internal and external variables. The majority of these specifications are connected to design. When a system's capabilities are investigated, use cases are created, and actors are identified as a result. Use cases are nothing more than a logical list of the system's capability. Fig 3.1 shows the use case diagram of our model.



**Fig 3.1-Use case diagram**

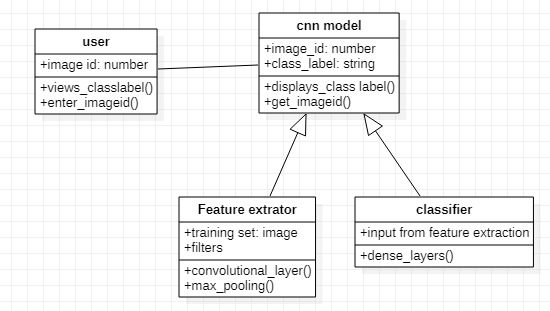
* 1. **Activity diagram**
  2. A flowchart depicting the flow of information from one action to the next is known as an activity diagram. A system operation can be used to describe the action. The control flow is shown from one action to the next. This flow can be sequential, branching, or simultaneously running. In activity diagrams, various elements such as fork, join, and others are used to deal with various types of flow control. Our model's activity diagram is shown in Figure -3.2.



**Fig. 3.2. Activity diagram**

**Class diagram:**

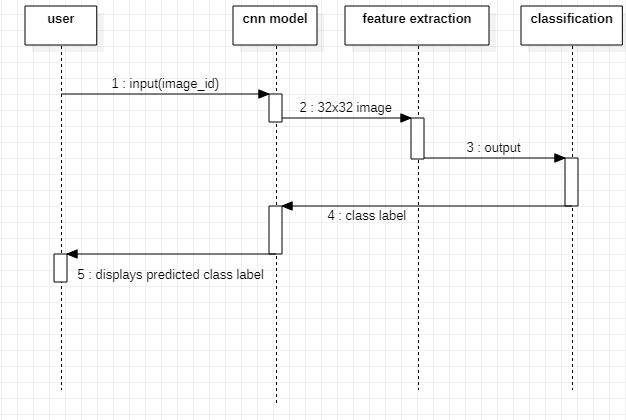
A static diagram is also referred to as a class diagram. It depicts an application's static view.A class diagram can be used to visualise, describe, and document many aspects of a system, as well as to create executable code for a software programme. The attributes and operations of a class, as well as the system's limitations, are depicted in a class diagram.Fig 3.3- shows the class diagram of our model.



**Fig. 3.3. Class diagram**

**Sequence diagram:**

A sequence diagram is a type of interaction diagram that shows how and in what order a group of items interacts. The order in which things interact, or the order in which these interactions occur, is depicted in a sequence diagram. Fig 3.4 shows Sequence diagram of our model.

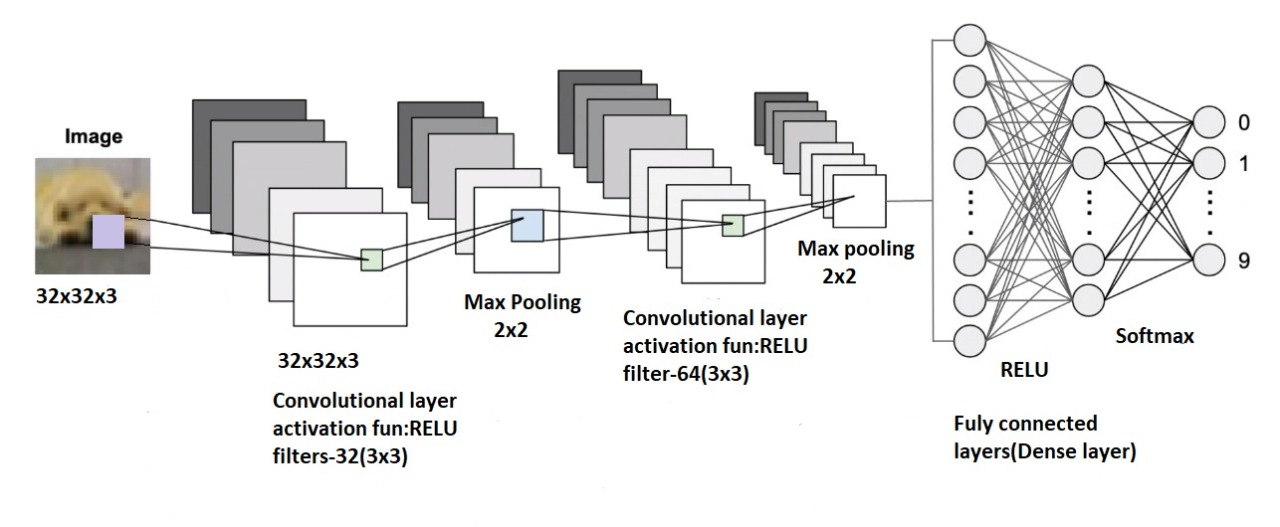


**Fig.3.4. Sequence diagram**

**System architecture**

A system architecture is a conceptual model that describes the structure, behaviour, and other characteristics of a system. An architecture description is a formal description and representation of a system that is organised in a way that makes reasoning about the system's structures and behaviors easier.

An architecture diagram is a system diagram that abstracts a software system's overall structure as well as the interactions, constraints, and boundaries between its components. It's an important tool because it provides a bird's-eye view of the software system's physical deployment and evolution path.Fig-3.5 shows the system architecture of our model.



**Fig. 3.5. System architecture**

**Chapter-4**

**CONSTRUCTION**

* 1. **Implementation**
  2. The project is implemented using the Python programming language.

To be more specific, Anaconda is being utilised for machine learning purposes. Anaconda is the most widely used Data Science platform for Data Scientists and IT professionals throughout the world.

Anaconda is a Python(and R) distribution. It is open-source and free, and it simplifies package management and distribution.

When we install Anaconda, we also install Jupyter, pandas, NumPy, Matplotlib, and scikit-learn, which are some of the most often used data science and machine learning tools.

**PYTHON**



* Python is a high-level scripting language that is interpreted, interactive, and object-oriented
* it is designed to be very readable. It frequently replaces punctuation with English keywords and has fewer syntactical constructions than other languages.

Python is a must-have skill for students and working professionals who want to become great software engineers, particularly in the Web Development field.

**ANACONDA**

 ****

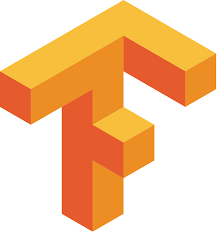
We observed Anaconda to be easier in Python technology due of the following factors:

1. Anaconda's package manager, conda, is a package manager as well as a virtual environment manager.

2. Conda does a better job of installing packages that need the user to compile C modules. This is significant for Microsoft Windows users, but not so much for GNU/Linux users.

3. It provides a Anaconda installation that contains only Python, conda, and a few other basic items, allowing you to install only the programs you require, either in the default environment or in a virtual environment.

**TENSORFLOW**

****

TensorFlow is an open-source machine learning platform that automates the entire process. It's a vast and adaptable ecosystem of tools, libraries, and other resources that provide workflow APIs at the highest level. The framework offers various layers of concepts from which to build and deploy machine learning models.

Some of the salient features are described below:

* Easy model building: TensorFlow allows you to construct and train models at different levels of abstraction.
* TensorFlow allows you to easily train and deploy your model regardless of the language or platform on which you're working
* Extensive experimentation for research: TensorFlow's Keras Functional API and Model Subclassing API provide flexibility and control for the development of complex topologies

**KERAS.**



TensorFlow is an open-source machine learning framework, while Keras is a neural network library that may be used for a variety of tasks. TensorFlow has high-level and low-level APIs, whereas Keras only has high-level APIs. Both frameworks offer high-level APIs for quickly developing and training models. Keras is a high-level neural network library built on top of TensorFlow, CNTK, and Theano. Keras is a deep learning framework that runs on both CPU and GPU and allows for quick prototyping. This framework is written in Python, making it simple to debug and extend. The main Keras advantages are listed below.

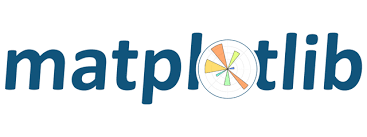
* User-friendly: Keras has a simple, consistent interface that is tailored to common use scenarios and provides clear, actionable feedback in the event of a user error.
* Modular and composable: Keras models are modular and composable, as they are constructed by assembling configurable building components with few constraints.
* Extendable: Keras allows you to quickly create custom building blocks for new ideas and research.
* Simple to use: Keras provides consistent and straightforward APIs that reduce the number of user activities required for common use cases while still providing clear and responsive feedback when a user error occurs.

**NUMPY**



NumPy is a Python library for interacting with arrays. It also includes matrices, Fourier transforms, and linear algebra functions. NumPy was created by Travis Oliphant in 2005. You are free to use it because it is an open source project. NumPy is the abbreviation for Numerical Python. NumPy functions that we used in our project include: creating a NumPy array, reshaping array dimensions, finding out the length of the NumPy array and slicing NumPy arrays.

**MATPLOTLIB**

****

Matplotlib is a visualisation tool that makes use of a Python-based low-level graph plotting library.John D. Hunter is the creator of Matplotlib. It is free and open source software that we can use. For platform compatibility, The majority of Matplotlib is written in Python, with a few exceptions in C, Objective-C, and JavaScript.

Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton gathered the Canadian Institute for Advanced Research (CIFAR-10) dataset. The CIFAR-10 dataset contains 60,000 32x32 color pictures divided into ten classes, each with 6,000 pictures. There are 50,000 photos for training and 10,000 photos for testing. There are five training batches and one test batch for each of the 10,000 photos in the dataset. The test batch consists of 1,000 photographs chosen at random from each class. The remaining photographs are randomly distributed in training batches, with some batches containing more photos from one class than others. Each of the training batches has exactly 5,000 photographs from each lesson.

**PROPOSED MODULES**

* Importing Packages:

We will import TensorFlow, matplotlib, keras, sklearn and NumPy.

* Load datasets:

we will load small images CIFAR-10 dataset from TensorFlow keras datasets.

* Feature extraction using CNN:

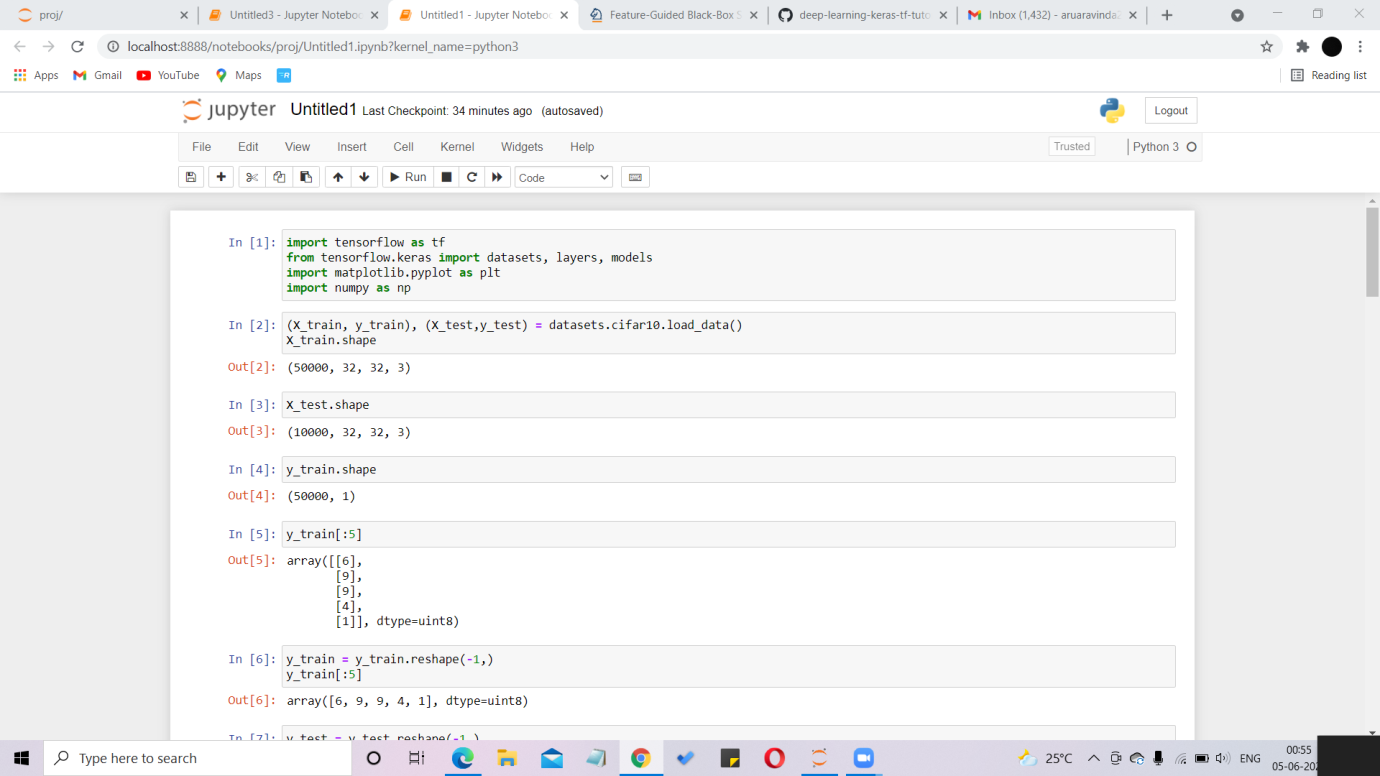
we perform feature extraction using convolution and max pooling layers.

* Classification:

Classification is done using Flattening and dense layers.

* Compile
* Evaluate
* Prediction

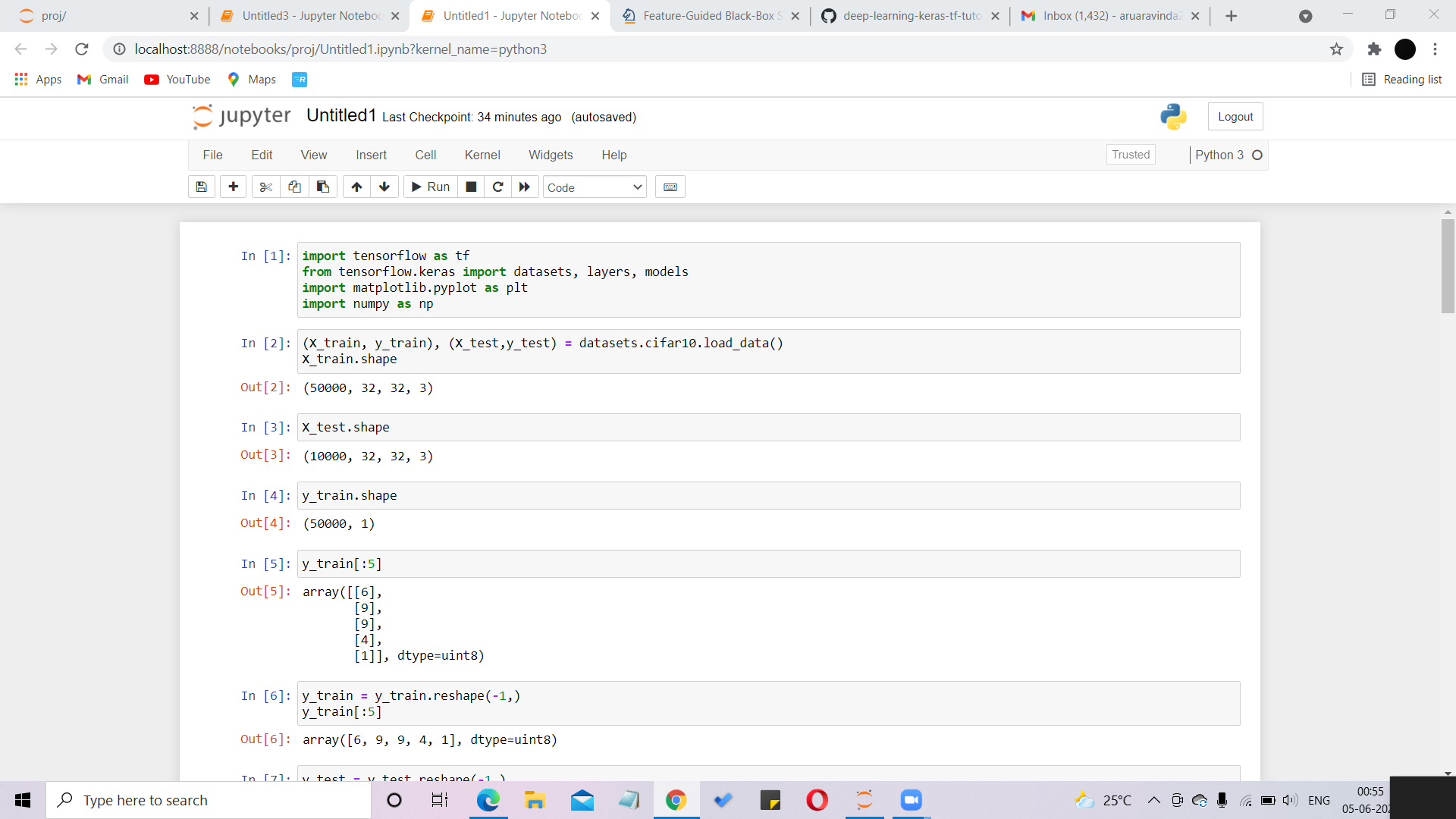
IMPORTING PACKAGES



**Fig4.1**

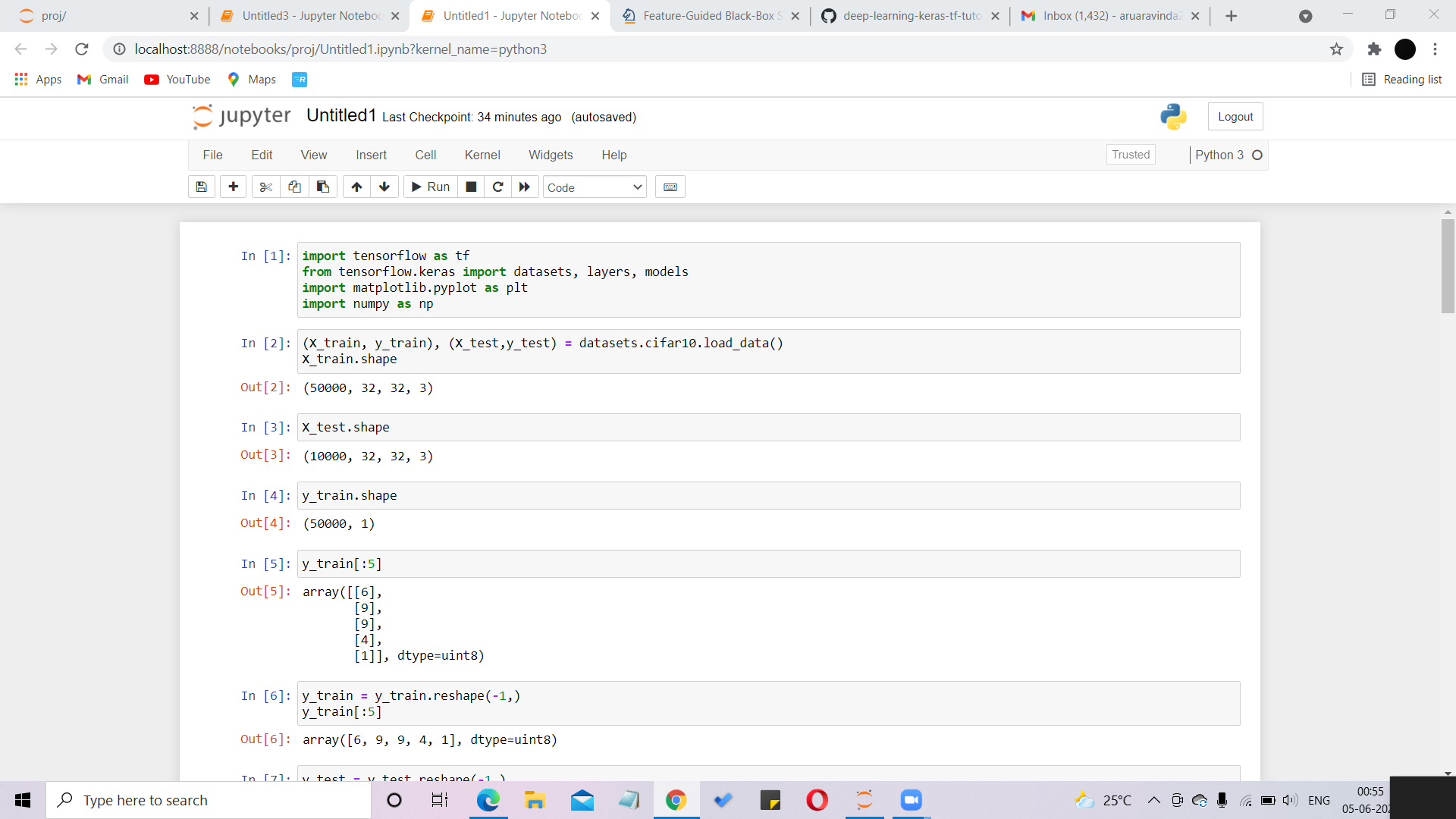
To build, train and predict our model, we use TensorFlow and keras, that can run on your Jupyter Notebook. we also imported NumPy and matplotlib for computing math functions, to plot graphs etc.

LOADING DATASET



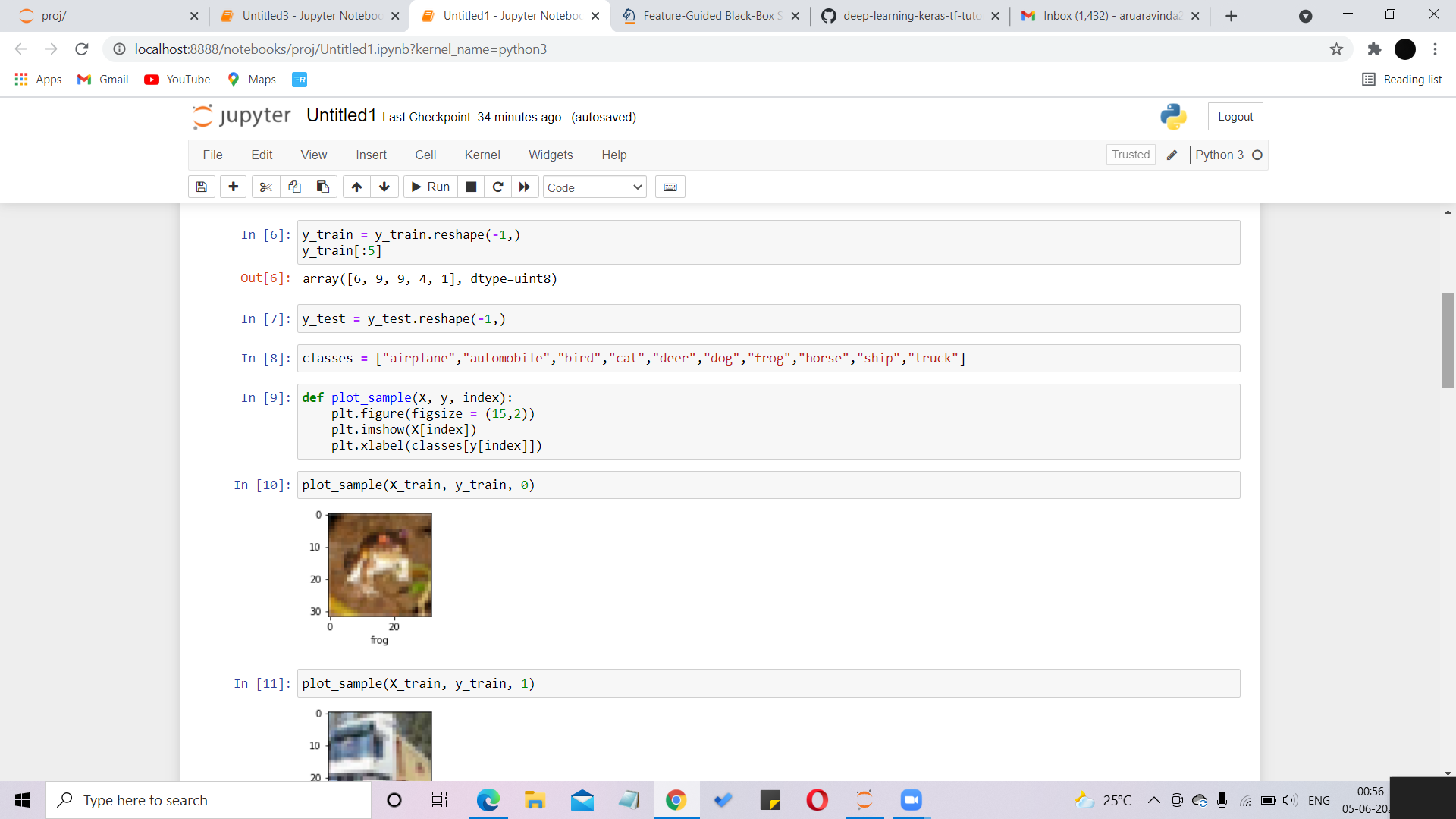
**Fig4.2**

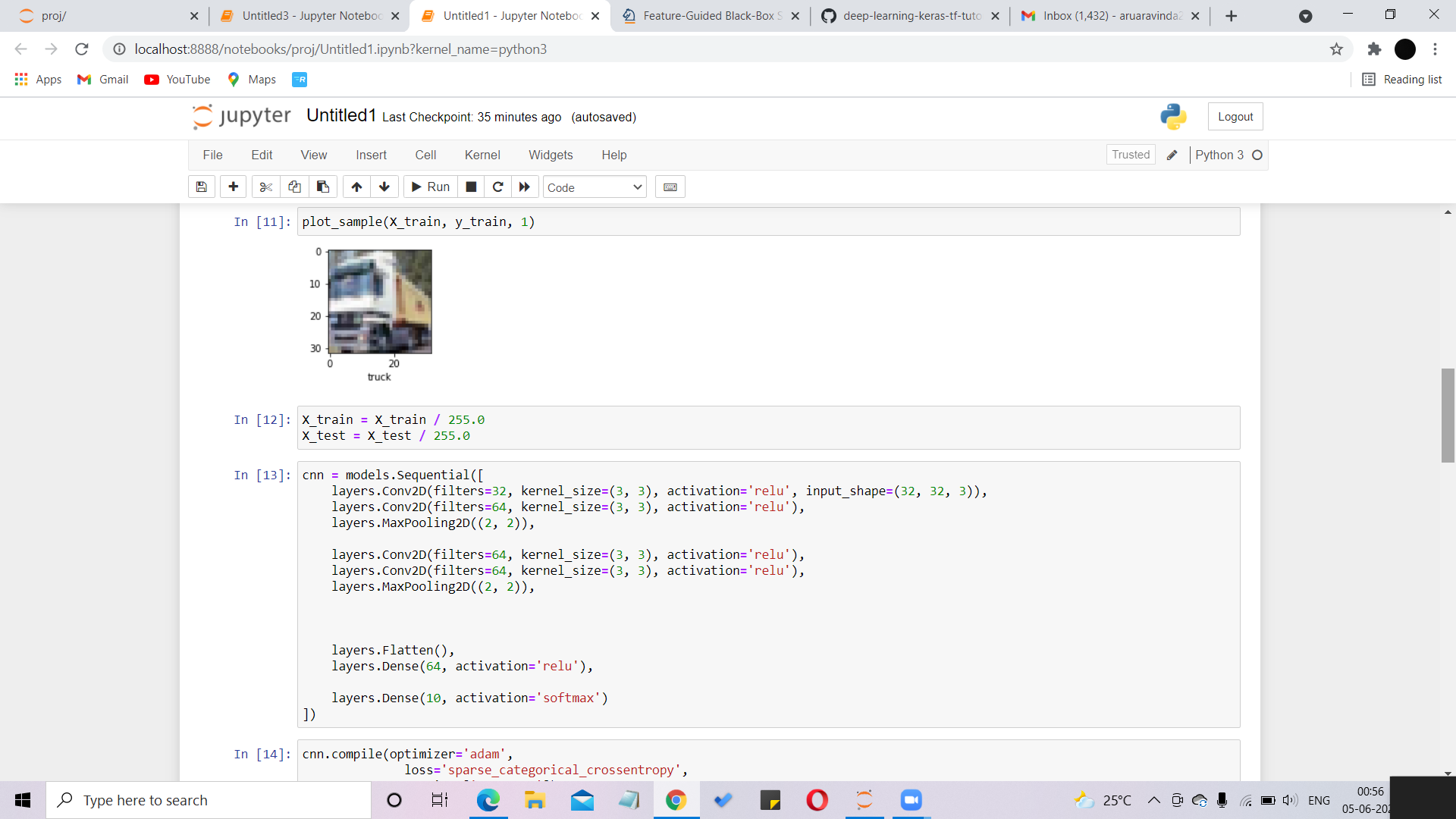
Here we use CIFAR 10 dataset in our model. In this dataset we have 50,000 images which are small and colored. They are of the size 32x32 pixels. We use 40,000 images for testing purpose and 10,000 images for validation purpose.

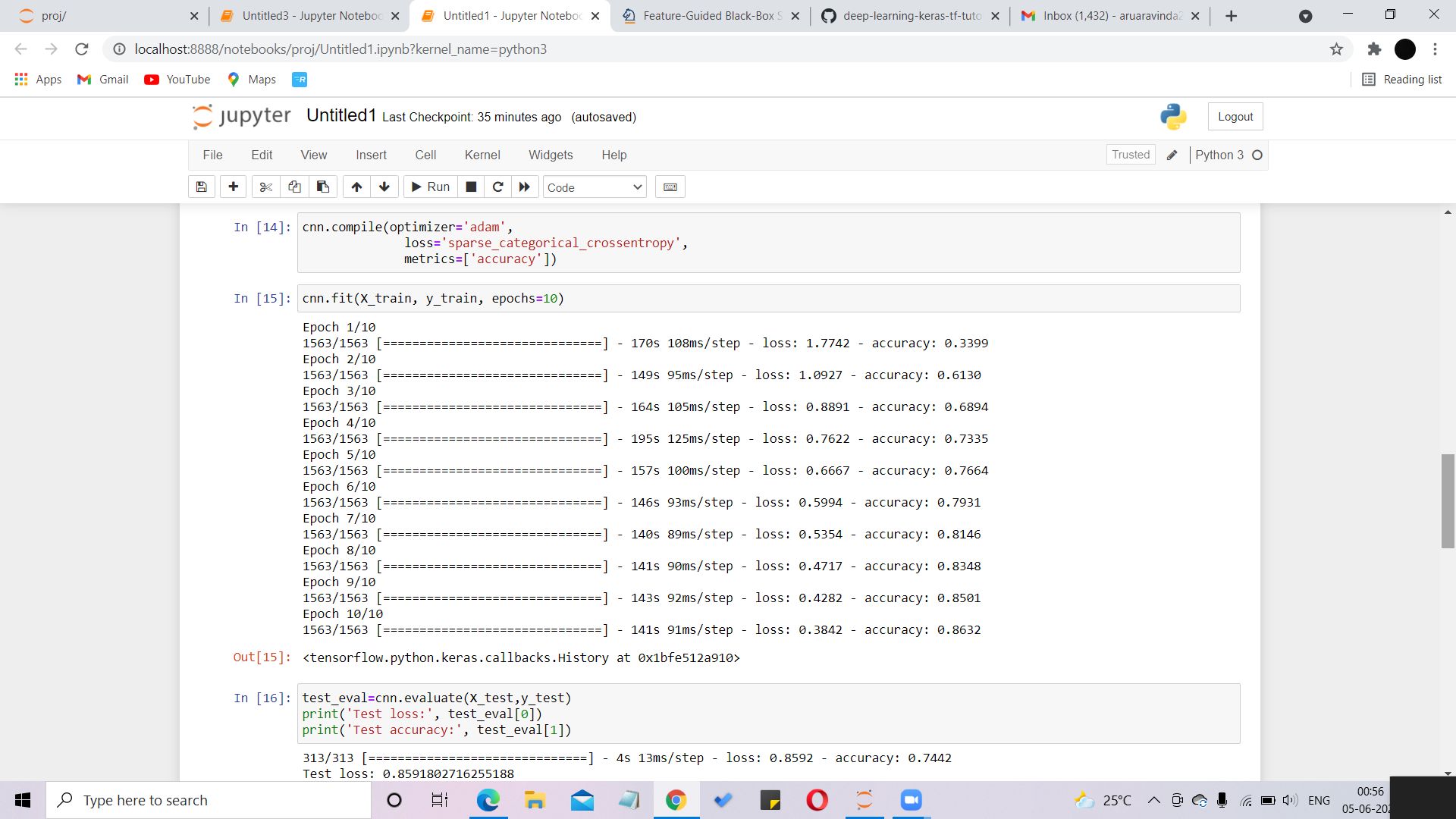


**Fig 4.3**

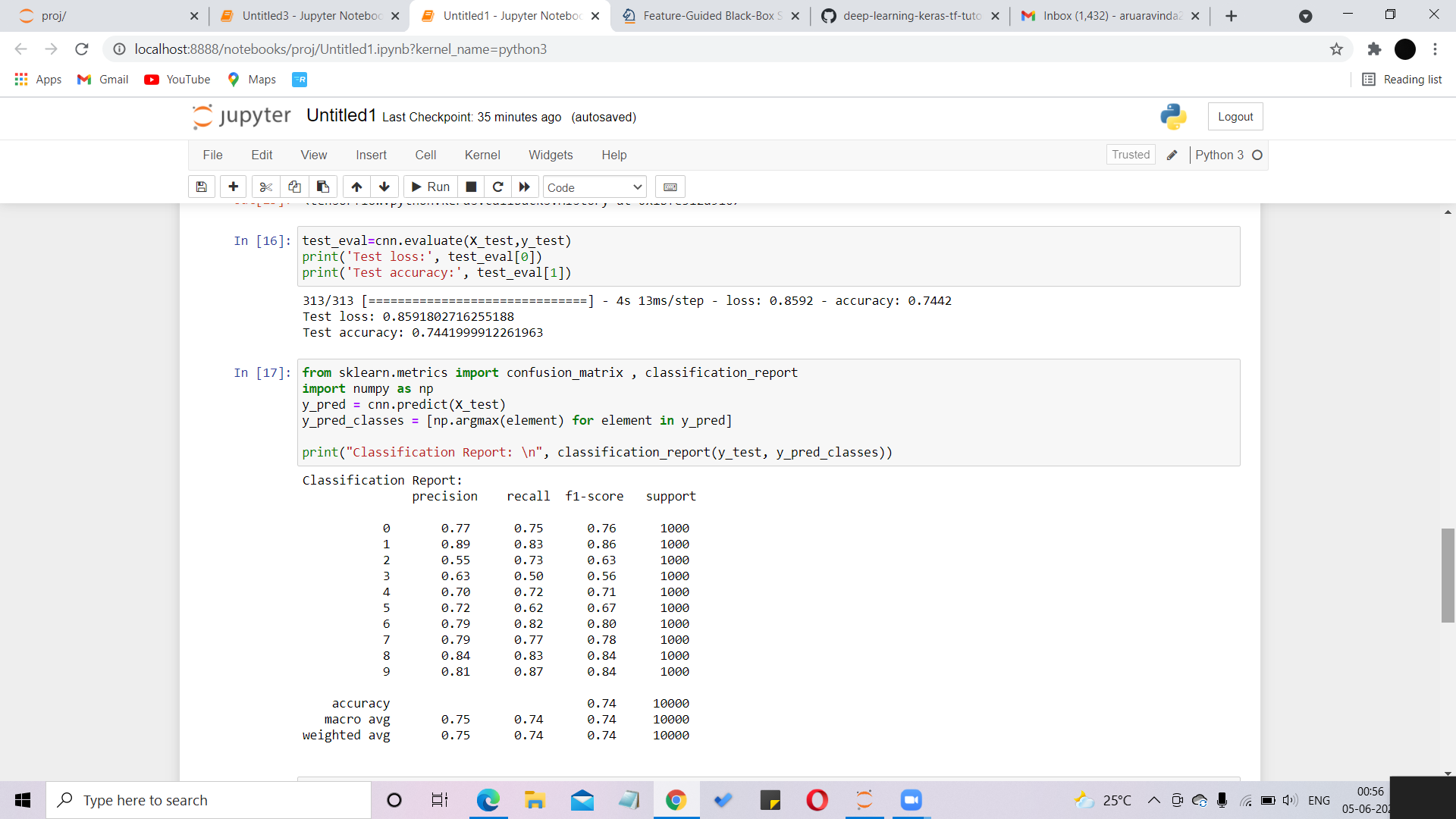
x set contains the test or validation dataset which contains 10,000 images and y set has the classes ranges from 0 to 9 and corresponds to the different classes of the CIFAR 10 dataset.



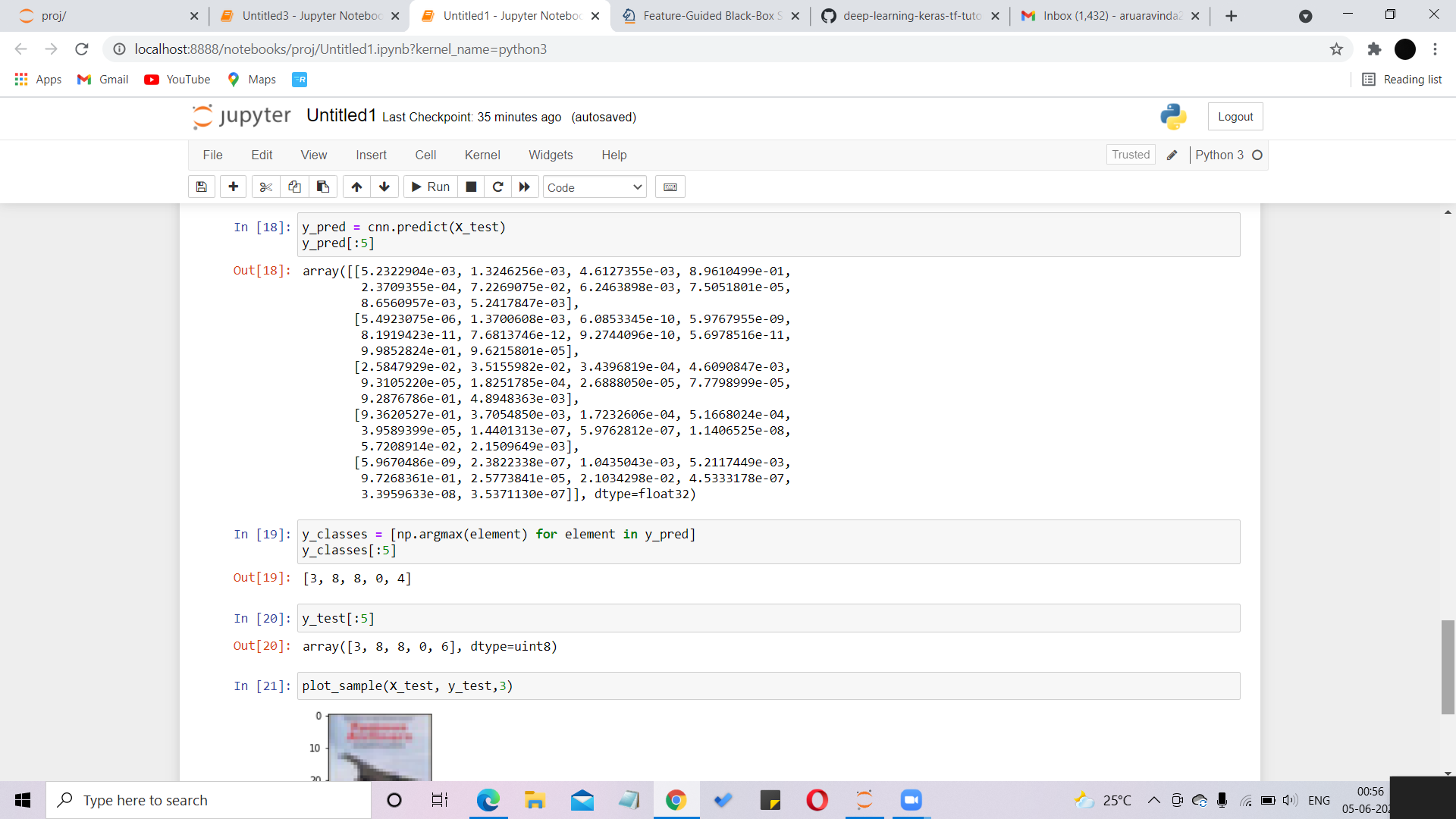
**Fig 4.4**

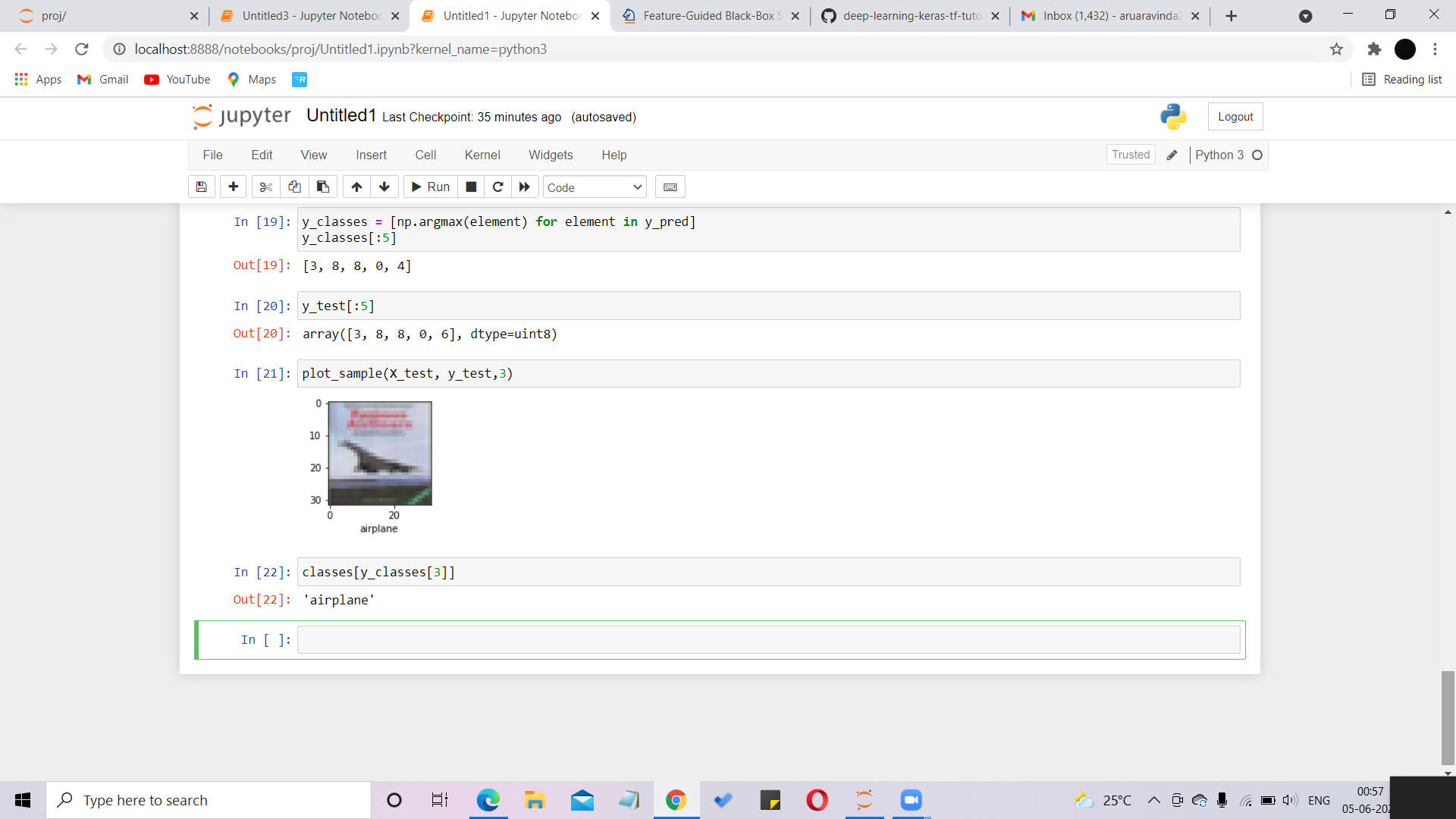
**Fig 4.5**

**Fig 4.6**



**Fig4.7**

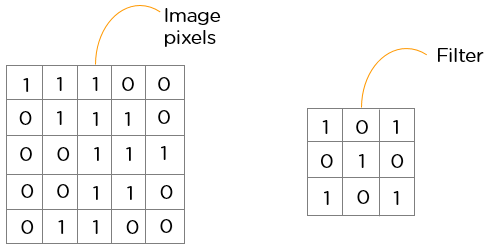


**Fig 4.8**

**Fig 4.9**

**Implementation Details**

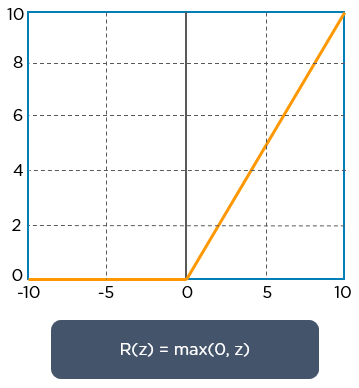
* **Convolution:** The image is inspected by filters that can segment the input image into smaller parts and generate feature maps (more precisely, it returns the number of feature maps equal to the number of selected filters). To get the result of this filtering method, we need a so-called trigger function. The resulting values ​​are mapped using trigger functions in the range 0 to 1, 1 to 1, etc. In this model, the ReLU function only sets a negative value to 0. The image is converted to image pixels and multiplied by a filter.Pixels and image filters, as shown in Figure 4.10.



**Fig 4.10**

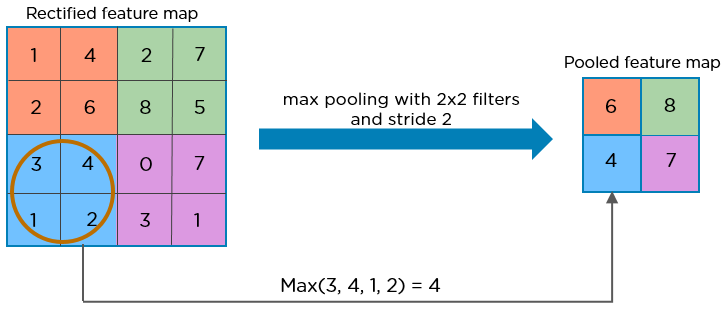
* **ReLu layer:**

The rectifier linear unit is once the feature maps have been extracted, we need to transfer them to the ReLU layer. ReLU sets all negative pixels to 0 and performs the operation item by item. This makes the network non-linear and results in a modified feature map. The ReLU function is shown in Figure 4.11. below.



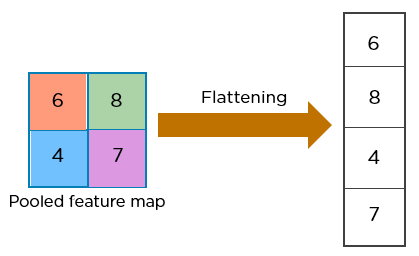
**Fig 4.11**

* **Pooling:** The main objective of this phase is to perform a function (in this example, the 'max' function returns the maximum value of a pixel) to reduce the size of the feature map. Here is an example of the maximum clustering function: Below Figure 4.12.



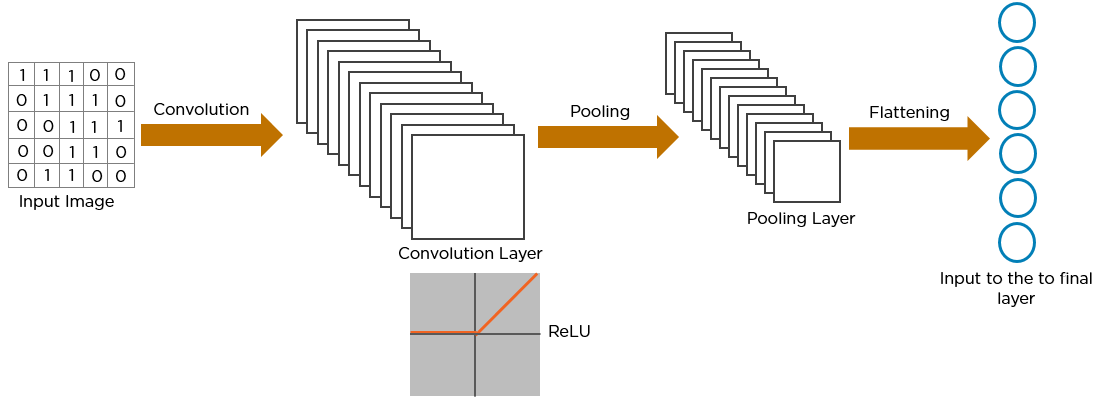
**Fig 4.12**

* **Flattening:** Flattening is the next step in the process. The 2D matrix obtained from the integrated feature map is flattened into a long, continuous linear vector. An example of flattening is shown in Figure 4.13 below.



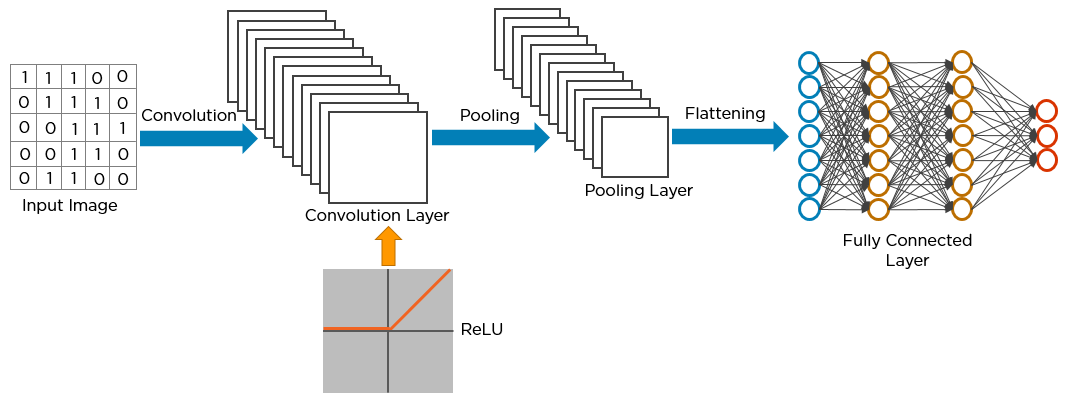
**Fig 4.13**

For image classification, a flattening matrix is ​​provided as input for the fully connected layers, as show in Figure 4.14 below.



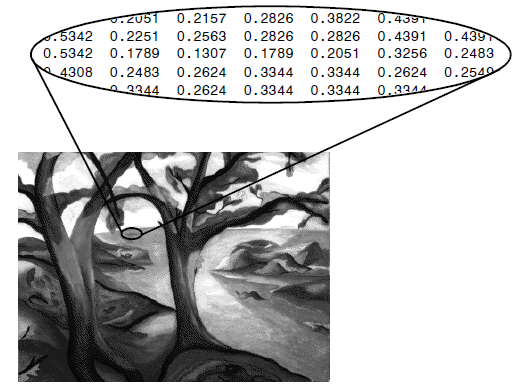
**Fig 4.14**

* **Fully connected:** The goal of the fully connected layer is to use these characteristics to classify the input images into different classes based on the training data set. After converting the raster image to a numeric matrix, reapply the trigger function to get the probability vector as follows: The final output is the same length as the layer vector (see Figure 4.15). In fact, the "SoftMax" trigger function is used which converts the input into a probability range. The range is 0 to 1 and the sum is 0 to 1. This method returns the probability of each class. In this case, the target class is the most probable class because it is a multi-classification task.

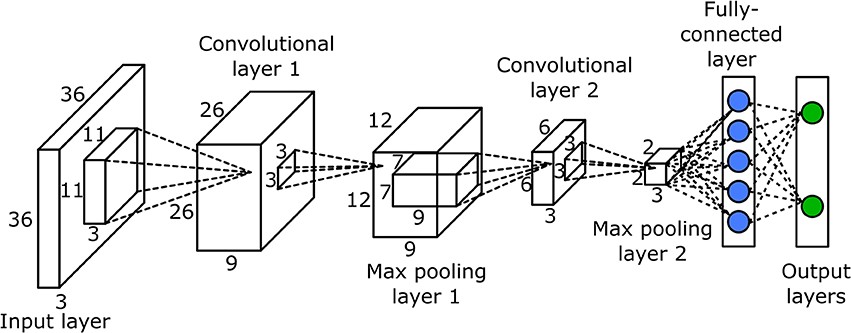


**Fig 4.15**

* **Loss function:** It assesses the model's consistency. If the conforming value deviates from the actual value, it returns higher values. Mean Squared Errors is a common loss function used in Linear Regression models (MSE). We'll utilize the Categorical Cross entropy function in this situation.
* **Optimizer:** The partial derivative of the loss function (called gradient) is determined with respect to the weights using an optimization function, and the weights are updated in the reverse direction of the obtained gradient. This cycle is repeated until the minimum value of the function reaches the loss. We'll use Adam Optimizer as an example.



**Fig. 4.16. pixels of an image.**



**Fig. 4.17. Layers in CNN.**

## Software Details

## • Anaconda Distribution (v5.1).

## • Python (3.6.5).

## • Jupyter Notebook.

## Hardware Details

## • Operating system: Windows 7 or newer, 64-bit macOS 10.9+, or Linux.

## • System architecture: 64-bit x86, 32-bit x86 with Windows or Linux.

## • CPU: Intel Core 2 Quad CPU Q6600 @ 2.40GHz or greater.

## • RAM: 4 GB or greater.

## Testing

Testing is the process of evaluating a system or its components to determine if they meet defined criteria. testing is the process of running your system to find vulnerabilities, bugs, or missing needs. T Testing is a consultation designed to provide interested parties with information on the quality of the software product or service being tested.

The program tests provide an objective and unbiased overview of the software to understand and understand the risks associated with implementing the software. The process of running a program or application to find errors and ensure that a software product is suitable for use is called testing.The test is the process of assessing the quality of one or more of interest through the execution of a software component or system. In general, these qualities reflect how well the component or system u

under test is. Meets the requirements of the design and development guide, is suitable for a wide range of inputs, performs its function in a reasonable amount of time, runs in the intended environment where it can be used and installed, and achieves the overall result desired by Produce.

## White box testing

White box testing is a software testing approach that involves testing the structure, architecture, and underlying code of a product to validate input and output flows and improve design, capabilities, and performance.. White box testing is also known as transparent box testing or open box testing. Test, test the transparent box, test the code base and test the glass box so that the code is visible on the tester. It is one of two components of software testing technique known as box testing. Black box testing, on the other hand, focuses on internal application behavior and internal software engineering testing.

## Black box testing

Black box testing is a software testing approach that involves testing the functionality of a software application without knowing the internal code structure, implementation details, or internal origins.

Black Box Testing is a type of software test that focuses on the inputs and outputs of software applications. Complete instructions on software requirements and specifications. Another name for this is crash test.

**Test cases:**

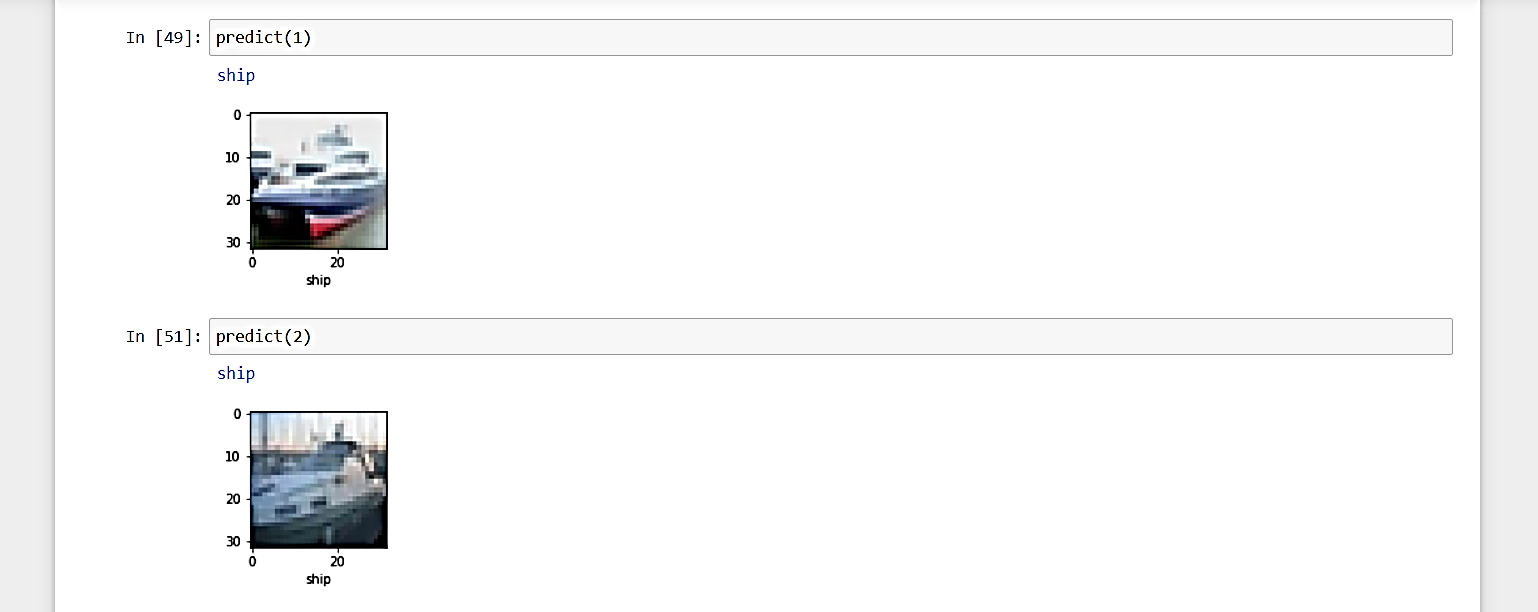
Here we define a function that plots the prediction vector and shows whether the predictions are correct (in blue) or incorrect (in red) with the image.

Since we are using 10,000 images for testing purpose, the index of the input can be in the range 0 to 9,999.

****

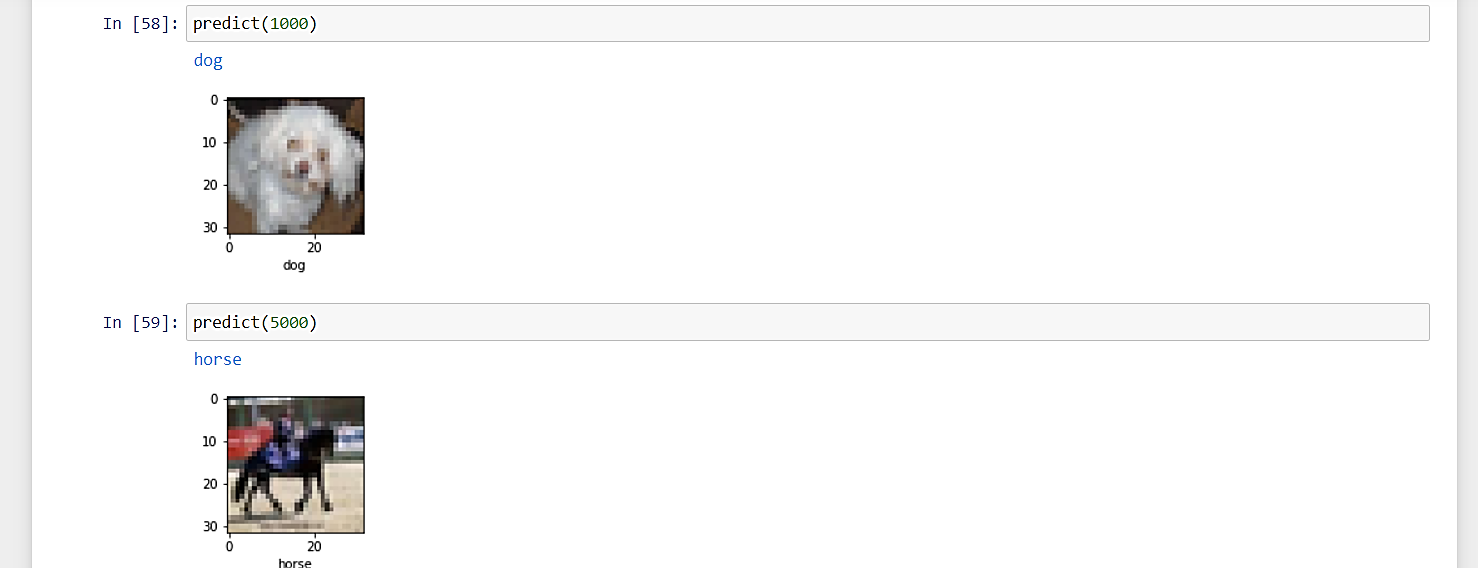
**Fig 4.18**

In this when we give the input value as 3 , we get its corresponding class name as airplane . This is done by layers which are convolutional layers, max pooling layers , fully connected layers etc. The corresponding image from the CIFAR 10 dataset is considered and then converted into its corresponding pixels values. Then the matrix of these pixels are then passed through the hidden layers of CNN model, which undergo certain operations through activation functions, filters, flattening etc. Based on these features are extracted and the image class is predicted by the model. This is verified with the original class label and the predicted class label is returned as blue if correct and red if wrong.



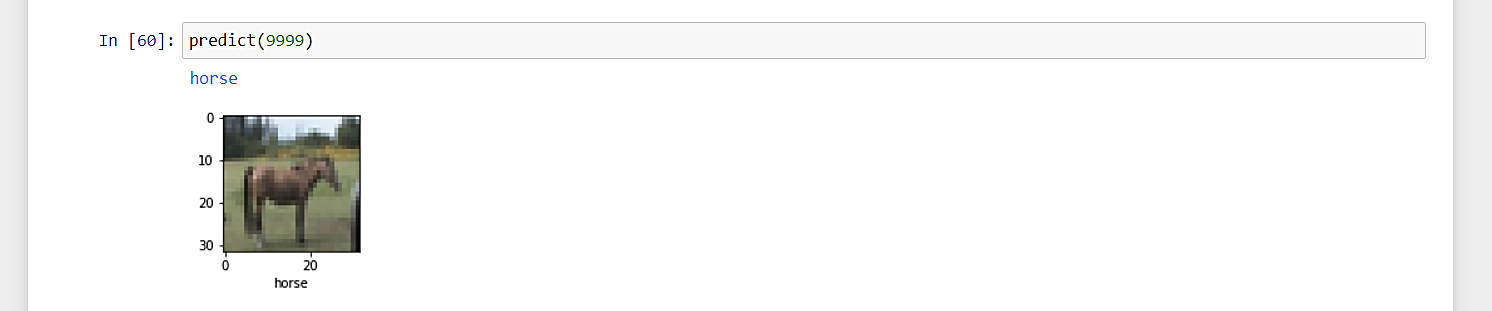
**Fig 4.19**

When we input the value as 1 and 2 to the predict function, we get the predicted class label as ship and ship in blue color which indicates that the predicted class is true respectively as shown in fig 4.19.



**Fig 4.20**

When we input the value as 1000 and 5000 to the predict function, we get the predicted class label as dog and horse in blue color which indicates that the predicted class is true respectively as shown in fig 4.20.



**Fig 4.21**

When we input the value as 9999 to the predict function, we get the predicted class label as horse in blue color which indicates that the predicted class is true as shown in fig 4.19.

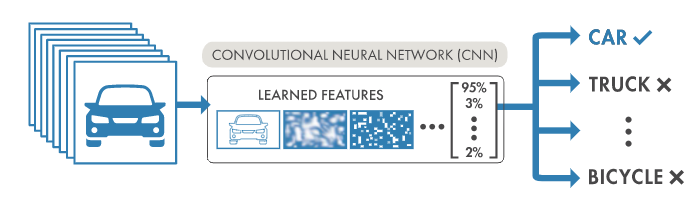
In this way, we can identify whether our model predicts the correct image class label or not. The predict function also returns the image with its original class label using matplotlib function which can be used as further verification process.

**Chapter-5**

**CONCLUSION AND FUTURE SCOPE**

* 1. **Conclusion**

Using photos from the CIFAR-10 data set, we used Convolutional Neural Networks (CNN) for image recognition. This data collection is utilized for training as well as testing. The graphics used for training purposes are small and colored. This dataset contains 60,000 pictures, each of which is an image, with 50,000 being used for training and 10,000 being used for testing. After the input image is passed through several layers we get the predicted class value as shown in fig 5.1 The most robust object categorization and detection algorithms now use deep learning architectures with numerous specialised layers for automating the filtering and feature extraction process. Linear regression, support vector machines, and decision trees are all machine learning algorithms with their own learning processes, but they all follow the same basic steps: Making a prediction, receiving a correction, and then adjusting the prediction mechanism depending on the correction is remarkably similar to how humans learn. Deep learning evolved as a novel solution to the problem, seeking to overcome past flaws by learning abstraction in data using a stratified description paradigm based on a nonlinear transformation. A fundamental advantage of deep learning is its capacity to do semi-supervised or unsupervised feature extraction over large datasets. We are currently living in the era of machine learning applications, and as these approaches are employed and enhanced, more will be developed in the next years. In the near future, significant advances in the development of even more specific hardware, as well as fundamental improvements in related mathematical theory, are expected, making artificial intelligence more present and crucial in today's world.

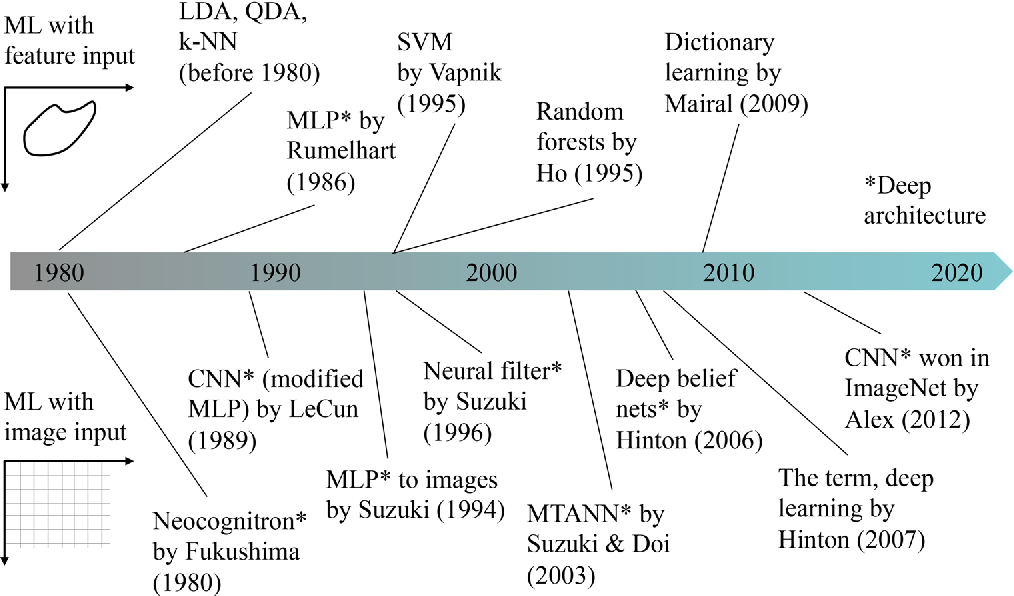


## Fig 5.1

## Future Scope

Because of its strong association with video analysis and picture understanding, object detection has gained a lot of academic attention in recent years (as shown in fig 5.2). Traditional object detection approaches are built on the basis of handcrafted features and shallow trainable architectures. Their performance may easily be improved by developing complicated ensembles that integrate several low-level image attributes with high-level context from object detectors and scene classifiers. More powerful tools that can learn semantic, high-level, and deeper features have been introduced with the rapid growth of deep learning to address the difficulties that plague older systems. These models react differently in terms of network architecture, training approach, and optimization function.

Although deep learning has made incredible strides recently, there are still obstacles to its implementation in medical imaging. Research on the susceptibility of deep neural networks in medical imaging is critical because, in comparison to non-medical tasks like recognising cats and dogs, clinical applications of deep learning require extreme robustness for ultimate usage in patients. It has a wide range of potential applications, including preventing traffic accidents, alerts of unsafe commodities in factories, military restricted area monitoring, and improved human–computer interface.



## Fig 5.2

Surveillance, face identification, fault detection, and character recognition are just a few of the uses for the object recognition system. The purpose of this thesis is to develop an object recognition system that can discriminate between two-dimensional and three-dimensional items in a picture. The performance of the object recognition system is determined by the features employed and the classifier used for recognition.

The goal of this research is to present a new feature extraction approach for extracting global and local features from the study area. Furthermore, the goal of the research is to combine traditional classifiers in order to recognise the item.

## Expected Outcome

The fundamental goal of our CNN model is to provide a prediction output that is as similar to the real one as possible. Furthermore, the algorithm can learn from its previous courses by re-weighting some parameters and minimising error terms once it has been evaluated.

We use TensorFlow, and one of the great things about TensorFlow and Keras is that you can create your own ‘homemade' CNN by adjusting the number and kind of layers based on the data you have. Your model will then use back propagation to tweak its parameter and minimise the error term on its own.

In this when we give an integer as the input value , we expect that the corresponding predicted class name is returned . Convolutional layers, max pooling layers, completely linked layers, and other layers help with this.. We expect that the model converts the corresponding image from the CIFAR-10 dataset into its corresponding pixels values. Then the matrix of these pixels are then passed through the hidden layers of CNN model, which undergo certain operations through activation funtions, filters, flattening etc.

Based on these we expect the model to extract features and predict the image class. This should be verified with the original class label and the predicted class label should be returned as blue if correct and red if wrong.Along with the image, the output of the function that plots our prediction vector displays whether the predicted value is correct (in blue) or incorrect (in red).

# References

[1]. DEEP LEARNING Book by Aaron Courvelle , Ian Goodfellow and Yoshua Bengio

[2]. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

[3]. Research papers by Elsevier.

[4]. <https://semiengineering.com/deep-learning-spreads/>

[5].<https://www.researchgate.net/publication/337146307_Convolutional_Neural_Networks>

[6]. <https://www.kaggle.com/c/cifar-10> to download the required dataset

[7].<https://www.simplilearn.com/tutorials/deep-learning-tutorial/convolutional-neural-network>

[8].<https://towardsdatascience.com/artificial-intelligence-vs-machine-learning-vs-deep-learning-2210ba8cc4ac>

[9]. <https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/?sh=54be0a768d4b>

[10].<https://www.kdnuggets.com/2016/10/artificial-intelligence-deep-learning-neural-networks-explained.html>