A REPORT OF ONE MONTH TRAINING

At

ANSH INFOTECH

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR

THE AWARD OF THE DEGREE OF

BACHELOR OF TECHNOLOGY

(Computer Science Engineering)

JULY 2023

SUBMITTED BY:

NAME – SIMARPREET KAUR

CLASS – CSE-C1

URN – 2104193, CRN - 2115137

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

GURU NANAK DEV ENGINEERING COLLEGE LUDHIANA

(An Autonomous College Under UGC ACT)

CERTIFICATE BY INSTITUTE

ANSH INFOTECH

DEVELOPMENT
TRAINING
CONSULTANCY
An ISO 9001-2015 Certified Company

AIT
Challenging the convention

Ref. No. AIT/IT/AI/2308/2751 Dated. 16 Aug 23

Certificate of Training

This is to certify that

Ms. Simarpreet Kaur D/O Mr. Arvinder Pal Singh

of GNDEC, Ludhiana

has successfully completed her Training in

"Artificial Intelligence"

From 30th June 2023 to 12th August 2023 at our organization.

During Training Her Performance was Excellent.

We wish her success for her future endeavor.

CERTIFIED ISO 9001:2015 COMPANY

Authorized Signature

SCF-4, 3rd Floor, Model Town Ext.,
D-Block Market, Dugri Road,
Ludhiana-141003.
E-mail.: anshinfotech1@gmail.com
Website: www.anshinfotech.org

M.: +91-84278-99400, 94171-68347

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA

CANDIDATE'S DECLARATION

We hereby declare that we have undertaken a technical report at "Guru Nanak Dev Engineering College" during a period from August to September in partial fulfilment of requirements for the award of degree of B.Tech (Computer Science Engineering) at Guru Nanak Dev Engineering College, Ludhiana. The work which is being presented in the training report submitted to Department of Computer Science Engineering at Guru Nanak Dev Engineering College, Ludhiana is an authentic record of training work.

Signature of Student

The four week industrial training Viva-Voce Examination of_____ has been held on_____ and accepted.

Signature of Internal Examiner Signature of External Examiner

ABSTRACT

Artificial intelligence (AI) is a rapidly growing field with the potential to revolutionize many aspects of our lives. AI systems are becoming increasingly capable of performing tasks that were once thought to be the exclusive domain of humans, such as driving cars, diagnosing diseases, and writing creative content.

This report provides an overview of the field of AI, its history, and its current state of development. It also discusses some of the potential benefits and risks of AI, and the challenges that need to be addressed in order to ensure that AI is used for good.

The report concludes by arguing that AI has the potential to be a powerful force for good in the world, but that it is important to use it responsibly and ethically.

Here are some of the key points that are covered in the report:

- The history of AI, from its early beginnings in the 1950s to its current state of rapid development.
- The different types of AI systems that are currently in use, such as machine learning, natural language processing, and computer vision.
- The potential benefits of AI, such as increased productivity, improved decision-making, and enhanced problem-solving.
- The risks of AI, such as job displacement, bias, and security.
- The challenges that need to be addressed in order to ensure that AI is used for good, such as developing ethical guidelines, ensuring that AI systems are not biased, and making AI systems secure from hacking.

The report concludes by arguing that AI has the potential to be a powerful force for good in the world, but that it is important to use it responsibly and ethically. It also calls for further research and development in AI, in order to ensure that this technology is used for the benefit of humanity.

ACKNOWLEDGEMENENT

I would like to express my deepest appreciation to all those who provided me the possibility to complete this report.

We would like to express our sincere gratitude to Ms. Gulabdeep Kaur, our project manager, for her guidance and support throughout this project. Her expertise in artificial intelligence was invaluable, and she helped us to stay on track and meet our deadlines. We would also like to thank the rest of our team for their hard work and dedication. This project would not have been possible without their contributions.

Ms. Kaur's expertise in Artificial Intelligence was essential to the success of this project. She helped us to understand the basics of Deep Learning, OpenCV and how to use it to implement our algorithms. She also provided us with valuable feedback on our work, and she helped us to identify and correct errors.

In addition to her technical expertise, Ms. Gulab was also a great team leader. She was always organized and efficient, and she kept us motivated throughout the project. She also created a positive and supportive work environment, which was essential for our success.

We are grateful for Ms. Gulab’s contributions to this project, and we are confident that she will continue to be a valuable asset to our team.

Thank you again, Ms. Gulab, for all of your help.

Sincerely,

Simarpreet Kaur

ABOUT THE INSTITUTE

Ansh Infotech is a leading institute that offers comprehensive professional skill development programs across a wide range of domains and cutting-edge technologies. With a diverse curriculum encompassing AI, ML, Data Science, Deep Learning, Mobile Application Development, Software Development, Full Stack/MERN Stack Development, Cyber Security, IoT, Robotics, Digital Marketing, Software Testing, Java, JS, Advance Java, and Hybrid Application Development, the institute equips learners with the expertise needed to excel in today's competitive landscape. What sets Ansh Infotech apart is its hands-on approach, providing students with valuable exposure to industrial projects, ensuring they gain practical experience alongside theoretical knowledge.

Ansh Infotech goes beyond education by bridging the gap between learning and industry demands. The institute offers Internships and JOB Opportunities to adopt candidates across various disciplines, empowering them to apply their skills in real-world scenarios. With a commitment to fostering talent and promoting innovation, Ansh Infotech stands as a catalyst for personal and professional growth, propelling individuals towards successful careers in the dynamic world of technology and beyond.

LIST OF FIGURES

3.1	Mode 1 and Mode 2 GUI Design
3.2	Mode 3 and Mode 4 GUI Design
3.3	Real Time Hand Recognition
3.4	Queue Visualisation
4.1	Building a faster self-checkout system

TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION

1.1 Understanding Artificial Intelligence: A Brief Overview

1.1.1 Delving into the Concept of Artificial Intelligence

Artificial Intelligence, often abbreviated as AI, represents a cutting-edge domain of computer science that aims to imbue machines with human-like intelligence. In essence, AI involves the development of algorithms and computational systems that enable machines to perceive, reason, learn, and act intelligently. It's not a mere buzzword; AI has evolved from its conceptual roots into a transformative force that underpins various sectors today.

1.1.2 Unveiling the Power of Machines that Think

The power of AI lies in its ability to replicate and exceed human cognitive functions. Machines equipped with AI algorithms can process vast amounts of data, identify patterns, and make predictions or decisions with remarkable accuracy. This capability has ushered in revolutionary changes across industries. From healthcare, where AI assists in early disease detection, to finance, where it optimizes investment portfolios, and even in autonomous vehicles, where it ensures safe navigation, the impact of AI is profound.

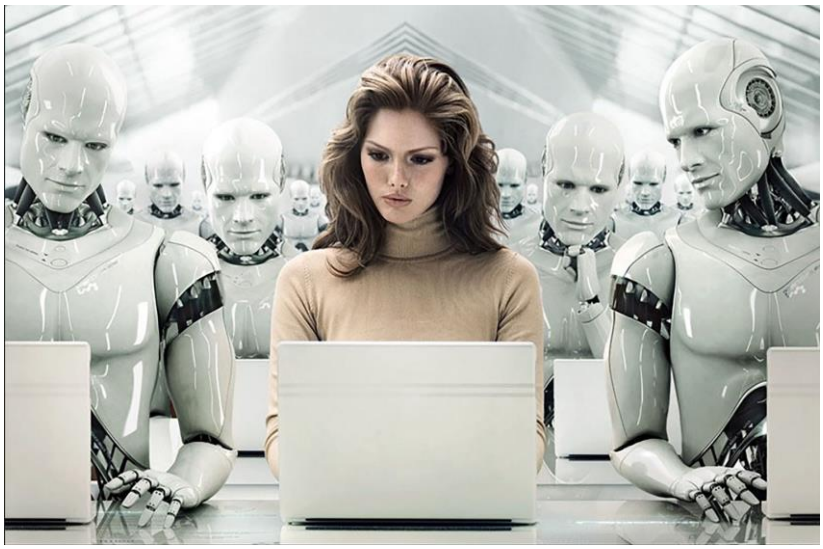


Fig-1.1 - *Human among robots*

1.2 Theoretical Foundations of Artificial Intelligence

1.2.1 Exploring the Core Principles of AI

At the heart of AI are foundational principles such as knowledge representation, problem-solving, and decision-making. Knowledge representation involves structuring information in a way that machines can interpret and use. Problem-solving encompasses devising algorithms that enable AI systems to find solutions to complex problems. Decision-making focuses on how AI systems choose the best course of action based on available data. These core principles provide the framework for AI development.

1.2.2 Cognitive Computing and Machine Learning: Building Blocks of AI

Cognitive computing and machine learning are fundamental to AI. Cognitive computing strives to emulate human thought processes, enabling machines to understand and respond to natural language and unstructured data. Machine learning, on the other hand, is a subset of AI that emphasizes the development of algorithms that allow systems to learn from data, improving their performance over time. These building blocks form the backbone of AI, driving advancements in various applications.

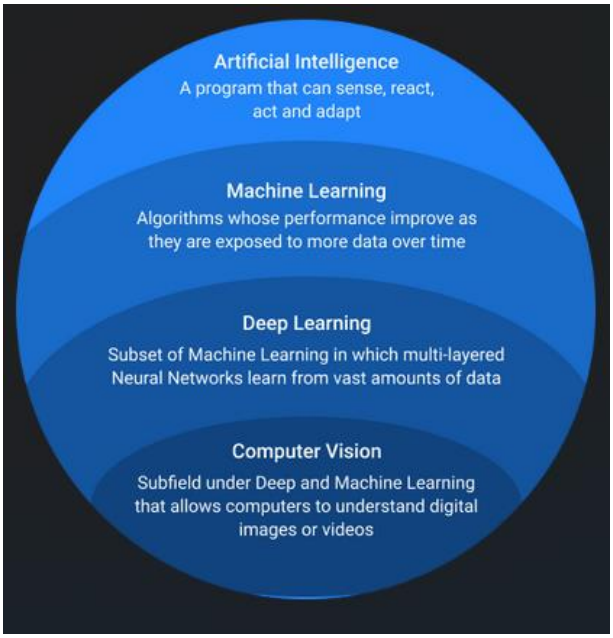


Fig-1.2- *Sub-fields of Artificial Intelligence*

1.3 Tools and Technologies in the World of AI

1.3.1 Navigating the AI Toolkit: Hardware and Software Essentials

In the AI landscape, a vast toolkit of hardware and software is available to developers. Popular software libraries like TensorFlow, scikit-learn, and PyTorch provide the necessary algorithms and tools for machine learning and deep learning tasks. Platforms like Google Colab and Jupyter Notebook offer interactive environments for AI model development. For deploying AI models, technologies like Flask enable the creation of web-based applications. The hardware side involves GPUs, TPUs, and even quantum computing for accelerating AI computations. Navigating this rich toolkit is crucial for effective AI development.

1.3.2 Python and Beyond: Programming Languages for AI Development

Python has emerged as the dominant programming language for AI due to its simplicity and a rich ecosystem of AI libraries. However, other languages like R and Julia also play essential roles in specific AI tasks. The choice of programming language depends on the project's requirements and the developer's expertise. Python, in particular, excels in data manipulation, making it a top choice for many AI practitioners.

Fig-1.3-*Python Libraries*

1.4 From Automation to Autonomy: AI's Impact on Industries

AI's transformative potential is evident across diverse industries. In business, AI is revolutionizing decision-making by providing data-driven insights, enhancing efficiency, and automating routine tasks. Healthcare benefits from AI by facilitating more accurate diagnoses, personalized treatment plans, and drug discovery. Transportation is on the verge of a revolution with AI-powered autonomous vehicles, promising safer and more efficient mobility. However, these advancements also raise ethical and societal questions, such as AI's impact on employment and the need for responsible AI development. Understanding these multifaceted impacts is essential as we delve deeper into the realm of Artificial Intelligence.

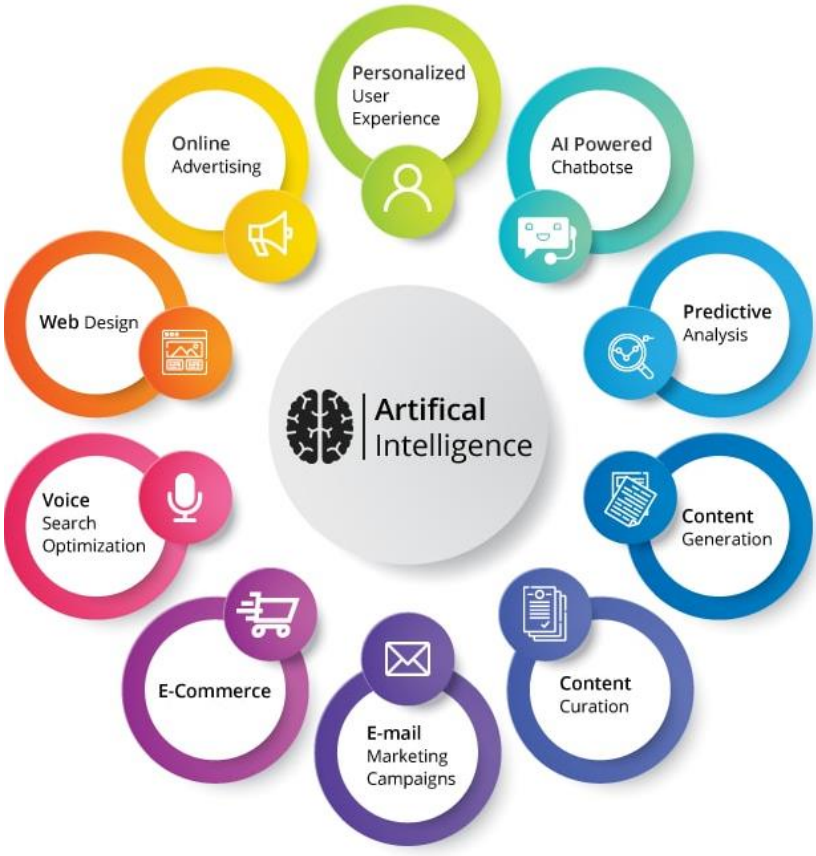


Fig-1.4- *Applications of Artificial Intelligence*

CHAPTER 2 – TRAINING WORK UNDERTAKEN

2.1 Building Foundations: Exploring Neural Networks and Deep Learning

2.1.1 Neural Network Basics

Neural networks are the foundation of deep learning, enabling machines to learn from data and make complex decisions. They are designed to mimic the interconnected structure of neurons in the human brain.

1. Neurons and Activation Functions:

Neurons are the building blocks of neural networks. Each neuron receives input, processes it, and produces an output. The core concepts include:

- **Input**: Neurons receive input data, often represented as feature vectors. Each input is associated with a weight that reflects its significance.

- **Weighted Sum:** Neurons compute a weighted sum of the inputs and their respective weights. This operation captures the impact of each input on the neuron's activation.

- **Bias:** A bias term is added to the weighted sum before passing it through an activation function. Bias allows for shifting the activation function horizontally.

- **Activation Function:** The weighted sum is passed through an activation function. Common activation functions include sigmoid, tanh, and ReLU (Rectified Linear Unit).

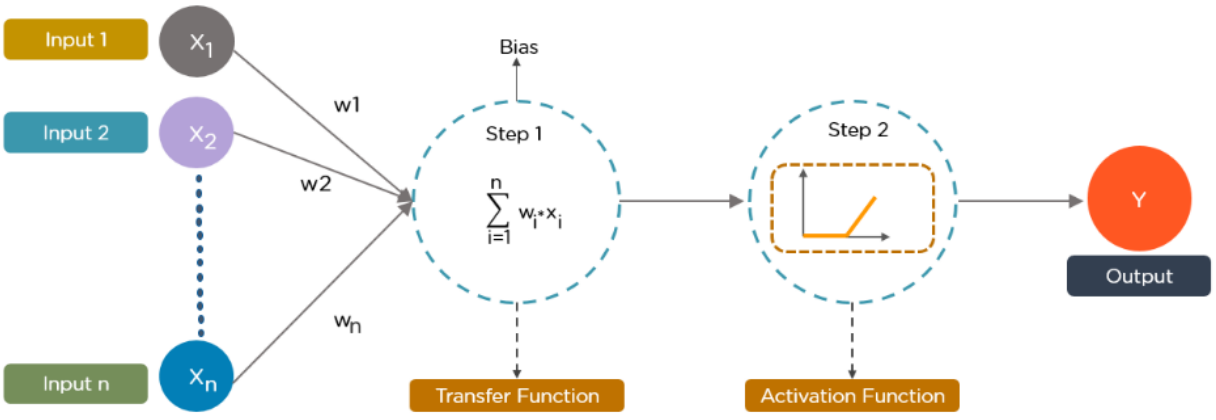


Fig-2.1- *Activation Function in Neural Networks*

2. Layers in a Neural Network:

A typical neural network consists of three main types of layers:

- Input Layer: This is the initial layer that receives the raw input data. The number of neurons in this layer corresponds to the number of features in the input data.

- Hidden Layers: Hidden layers are intermediary layers between the input and output layers. They perform computations and progressively extract higher-level features from the data.

- Output Layer: The output layer produces the final prediction or classification result. The number of neurons in the output layer depends on the problem type—regression, binary classification, or multiclass classification.

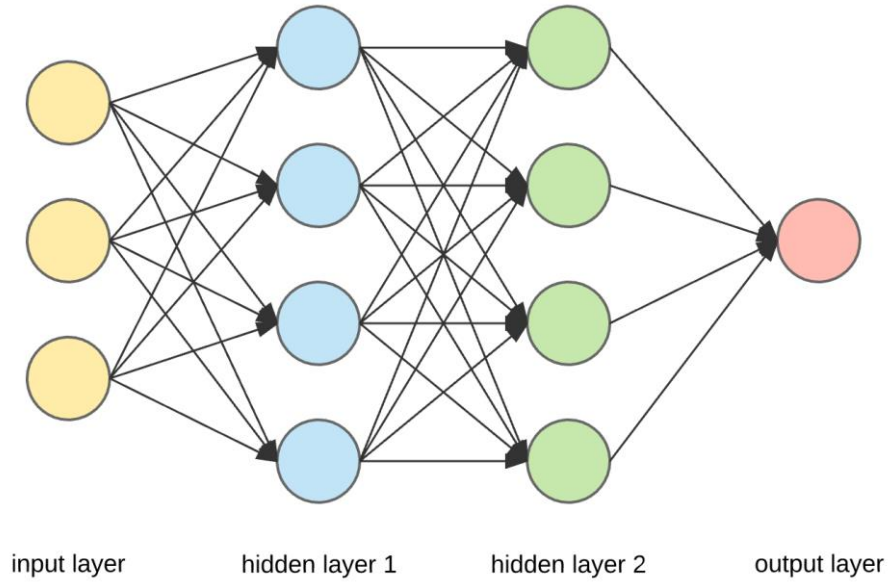


Fig-2.2- *Layers in Neural Networks*

3. Feedforward Process:

Neural networks operate using a feedforward process, where information flows from the input layer through the hidden layers to the output layer. The steps involved include:

- Input Propagation: Input data is fed into the input layer. Each neuron calculates its weighted sum and applies the activation function to produce an output.

- **Hidden Layer Computation:** Outputs from the previous layer serve as inputs to the next layer. This process is repeated across all hidden layers.

- **Output Generation:** The final hidden layer's outputs are used as inputs to the output layer. The output layer produces the model's prediction..

2.1.2 Deep Learning

2.1.2.1 Introduction to Deep Learning:

- Deep learning involves neural networks with multiple hidden layers, referred to as deep neural networks.

- Depth allows networks to learn hierarchical representations of data, capturing intricate patterns.

2.1.2.2 Artificial Neural Networks (ANNs)

The term "Artificial Neural Network" is derived from Biological neural networks that develop the structure of a human brain. Similar to the human brain that has neurons interconnected to one another, artificial neural networks also have neurons that are interconnected to one another in various layers of the networks. These neurons are known as nodes.

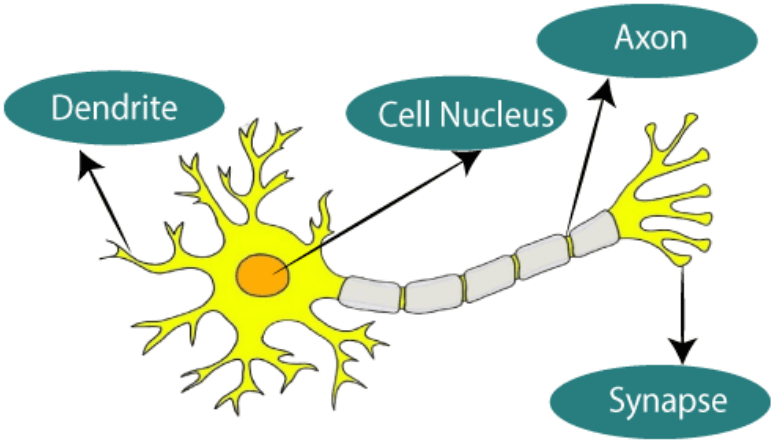


Fig-2.3- *Biological Neural Network or Structure of a Neuron*

Biological Neural Network	Artificial Neural Network
Dendrites	Inputs
Cell nucleus	Nodes
Synapse	Weights
Axon	Output

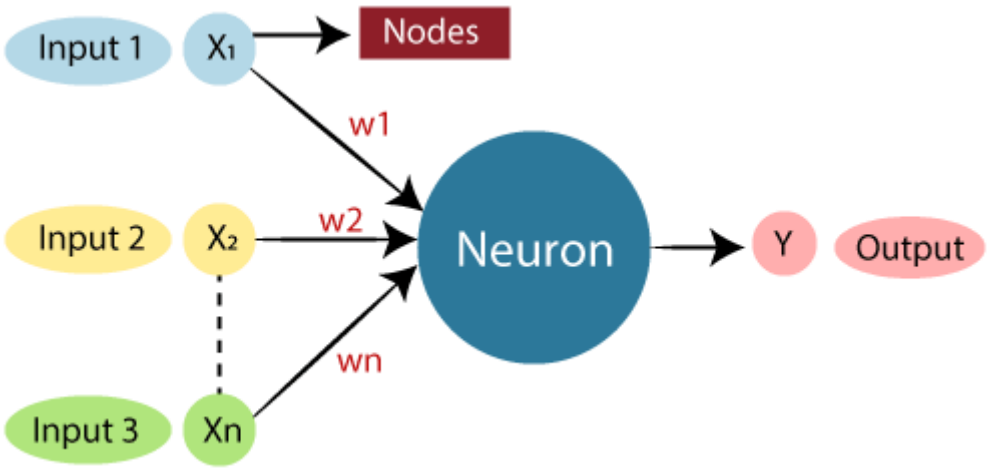


Fig-2.4- *Artificial Neural Network*

2.1.2.3 Architecture of ANN

- Input Layer: As the name suggests, it accepts inputs in several different formats provided by the programmer.

- Hidden Layer: The hidden layer presents in-between input and output layers. It performs all the calculations to find hidden features and patterns.

- Output Layer: The input goes through a series of transformations using the hidden layer, which finally results in output that is conveyed using this layer.

The artificial neural network takes input and computes the weighted sum of the inputs and includes a bias. This computation is represented in the form of a **transfer function.**

$$\sum_{i=1}^{n} Wi * Xi + b$$

It determines weighted total is passed as an input to an **activation function** to produce the output. **Activation functions** choose whether a node should fire or not.

2.2 Unleashing Visual Intelligence: Convolutional Neural Networks (CNN)

2.2.1 Introduction :

A **Convolutional Neural Network** (CNN) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with CNN. It uses a special technique called Convolution.

- Suited for image and spatial data, CNNs employ convolutional layers to extract features hierarchically.
- Max-pooling layers downsample the data, retaining relevant information.

2.2.2 Few Definitions :

1. Image Representation

Images have height and width, so the information contained in it is represented with a two dimensional structure (a matrix) but images have colors, and to add information about the colors, we need another dimension, and that is when Tensors become particularly helpful.

Images are encoded into color channels, the most common one being RGB, which means Red, Blue and Green.

Fig-2.5- *Representation of Image in 3 Channels in CNN*

2. Stride and Padding

Stride denotes **how many steps we are moving** in each steps in convolution. By default it is one.

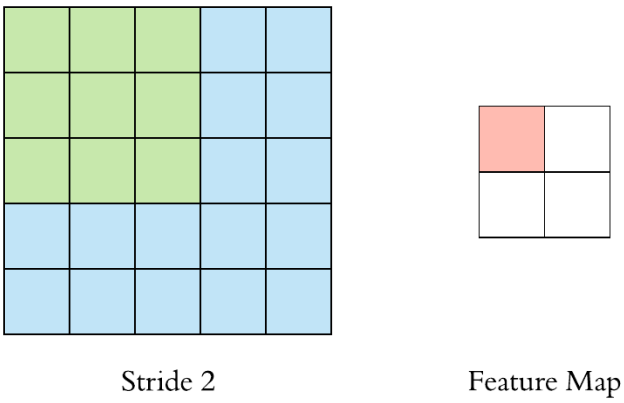


Fig-2.6- *Stride without Padding*

We can observe that the size of output is smaller that input. To maintain the dimension of output as in input , we use padding. **Padding** is a process of **adding zeros to the input matrix symmetrically**. In the following example,the extra grey blocks denote the padding. It is used to make the dimension of output same as input.

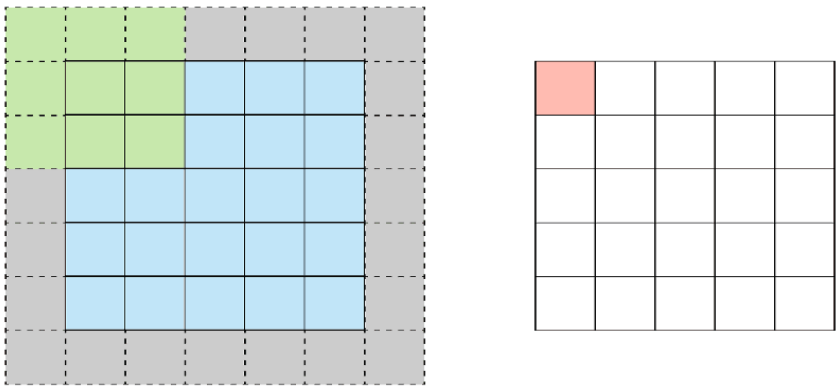


Fig-2.7- *Stride with Padding*

2.2.3 Architecture of CNN :

A CNN architecture is developed by a stack of different layers that convert the input volume into an output volume through a differentiable function. A few different types of layers are commonly used.

Below is the stack of different layers in CNNs:

- **Convolutional layers**
- **Pooling layer**
- **Fully connected layer**

In summary, the example of complete layers of CNNs:

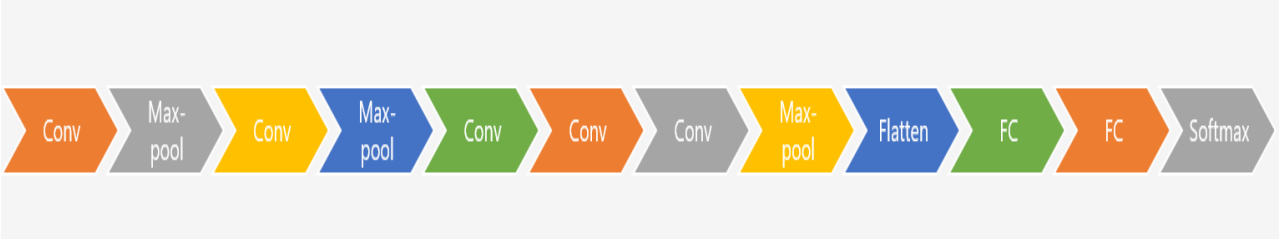


Fig-2.8 - *Complete Layers in CNN*

The complete architecture of CNNs:

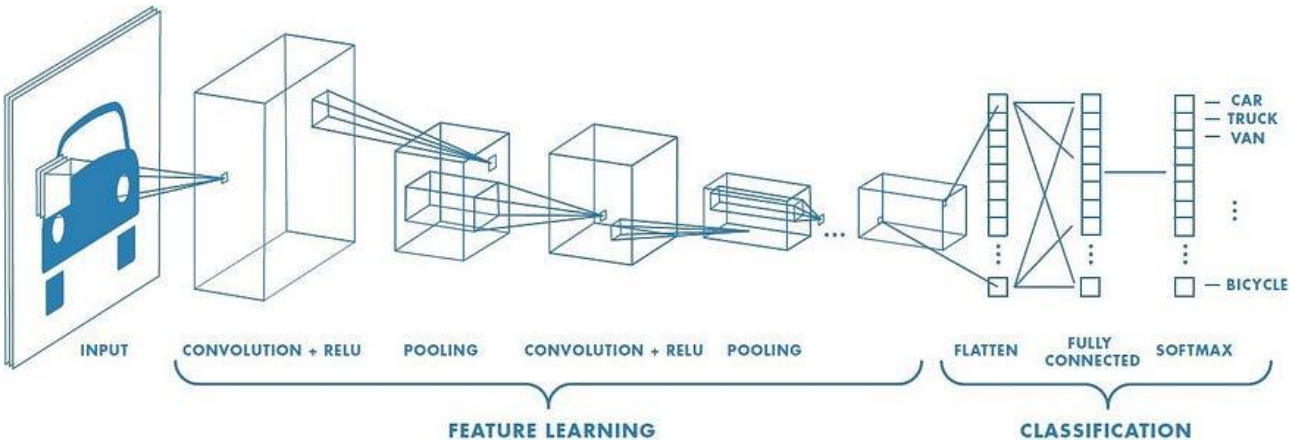


Fig-2.9- *Architecture of CNN*

- Input Layer : Input layer in CNN should contain image data.

- Convolutional layer : It is sometimes called feature extractor layer because features of the image are get extracted within this layer.The output will be the input for the next layer. Convolutional layer also contains ReLU activation to make all negative value to zero.

- Pooling layer : It is used to reduce the spatial volume of input image after convolution. It is used between two convolution layer. If we apply FC after Convo layer without applying pooling or max pooling, then it will be computationally expensive and we don't want it. So, the max pooling is only way to reduce the spatial volume of input image.

2.3 Textual Insights: Natural Language Processing (NLP)

2.3.1 Introduction :

NLP stands for Natural Language Processing, which is **a part of Computer Science, Human language,** and **Artificial Intelligence**. It is the technology that is used by machines to understand, analyse, manipulate, and interpret human's languages. It helps developers to organize knowledge for performing tasks such as **translation, automatic summarization.**

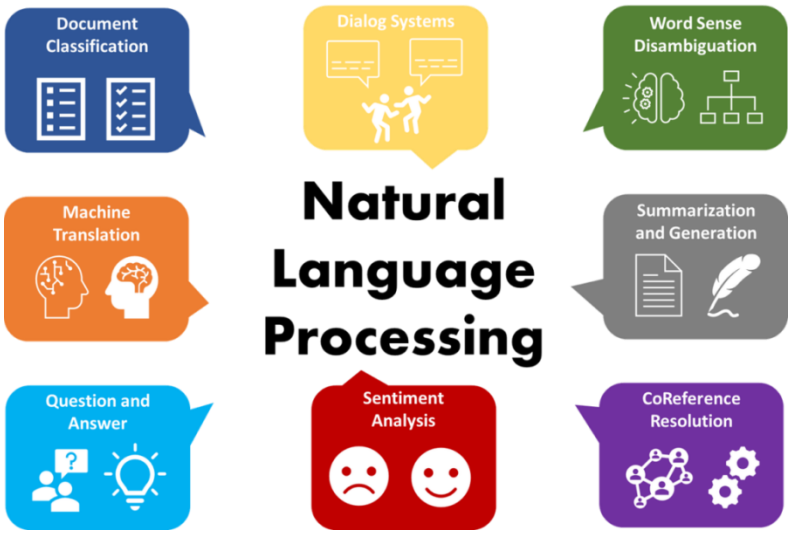


Fig-2.10- *Applications of NLP*

2.3.2 How to build an NLP pipeline :

There are the following steps to build an NLP pipeline -

Step1: Sentence Segmentation: Sentence Segment is the first step for building the NLP pipeline. It breaks the paragraph into separate sentences.

Step2: Word Tokenization: Word Tokenizer is used to break the sentence into separate words or tokens.

Step3: Stemming: Stemming is used to normalize words into its base form or root form. For example, celebrates, celebrated and celebrating, all these words are originated with a single root word "celebrate."

Step 4: Lemmatization: Lemmatization is quite similar to the Stamming.

Step 5: Identifying Stop Words: In English, there are a lot of words that appear very frequently like "is", "and", "the", and "a".

Step 6: Dependency Parsing: Dependency Parsing is used to find that how all the words in the sentence are related to each other.

Step 7: POS tags: POS stands for parts of speech, which includes Noun, verb, adverb, and Adjective. .

Step 8: Named Entity Recognition (NER): Named Entity Recognition (NER) is the process of detecting the named entity such as person name, movie name, organization name, or location.

Step 9: Chunking: Chunking is used to collect the individual piece of information and grouping them into bigger pieces of sentences.

Phases of NLP

There are the following five phases of NLP:

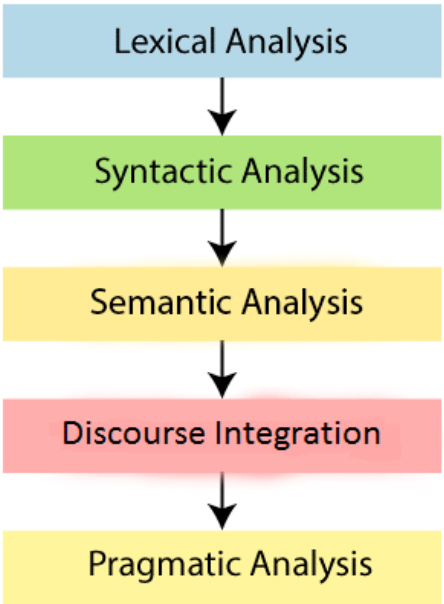


Fig-2.11- *Phases of Natural Language Processing*

1. Lexical Analysis and Morphological: The first phase of NLP is the Lexical Analysis. This phase scans the source code as a stream of characters and converts it into meaningful lexemes. It divides the whole text into paragraphs, sentences, and words.

2. Syntactic Analysis (Parsing): Syntactic Analysis is used to check grammar, word arrangements, and shows the relationship among the words.

3. Semantic Analysis: Semantic analysis is concerned with the meaning representation. It mainly focuses on the literal meaning of words, phrases, and sentences.

4. Discourse Integration: Discourse Integration depends upon the sentences that proceeds it and also invokes the meaning of the sentences that follow it.

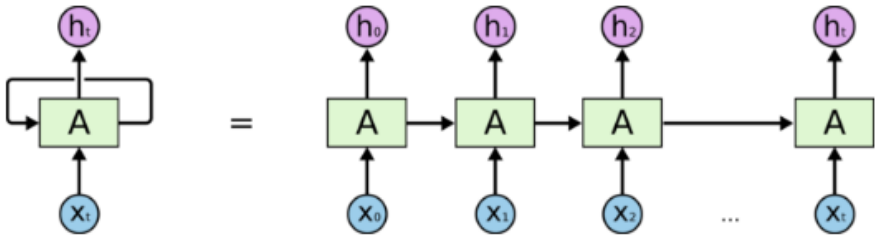
5. Pragmatic Analysis: Pragmatic is the fifth and last phase of NLP. It helps you to discover the intended effect by applying a set of rules that characterize cooperative dialogues.

2.4 Sequences and Context: Recurrent Neural Networks (RNN)

2.4.1 Introduction to RNN :

Recurrent Neural Network is a generalization of feedforward neural network that has an **internal memory**. RNN is recurrent in nature as it performs the same function for every input of data while the **output of the current input depends on the past one computation.**

Unlike feedforward neural networks, RNNs can use their **internal state (memory)** to process sequences of inputs.



An unrolled recurrent neural network.

First, it takes the X(0) from the sequence of input and then it outputs h(0) which together with X(1) is the input for the next step. So, the h(0) and X(1) is the input for the next step. Similarly,

h(1) from the next is the input with X(2) for the next step and so on. This way, it keeps remembering the context while training.

The formula for the current state is

$$h_t = f(h_{t-1}, x_t)$$

Applying Activation Function:

$$h_t = \tanh (W_{hh}h_{t-1} + W_{xh}x_t)$$

W is weight, h is the single hidden vector, Whh is the weight at previous hidden state, Whx is the weight at current input state, tanh is the Activation funtion, that implements a Non-linearity that squashes the activations to the range[-1.1]

Output: $y_t = W_{hy}h_t$

Yt is the output state. Why is the weight at the output state.

Advantages of Recurrent Neural Network

1. RNN can model sequence of data so that each sample can be assumed to be dependent on previous ones
2. Recurrent neural network are even used with convolutional layers to extend the effective pixel neighbourhood.

Disadvantages of Recurrent Neural Network

1. **Gradient vanishing** and **exploding** problems.
2. Training an RNN is a very difficult task.
3. It cannot process very long sequences if using tanh or relu as an activation function.

2.4.2 What is Long Short Term Memory (LSTM)?

Long Short-Term Memory (LSTM) networks are a modified version of recurrent neural networks, which makes it easier to remember past data in memory. **The vanishing gradient problem of RNN is resolved here.** LSTM is well-suited to classify, process and predict time series given time lags of unknown duration. It trains the model by using back-propagation. In an LSTM network, three gates are present:

LSTM gates

1. Input gate — discover which value from input should be used to modify the memory. Sigmoid function decides which values to let through 0,1. and tanh function gives weightage to the values which are passed deciding their level of importance ranging from-1 to 1.

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

2. Forget gate — discover what details to be discarded from the block. It is decided by the sigmoid function. it looks at the previous state(ht-1) and the content input(Xt) and outputs a number between 0(omit this)and 1(keep this)for each number in the cell state Ct−1.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

3. Output gate — the input and the memory of the block is used to decide the output. Sigmoid function decides which values to let through 0,1. and tanh function gives weightage to the values which are passed deciding their level of importance ranging from-1 to 1 and multiplied with output of Sigmoid.

$$o_t = \sigma\left(W_o \; [\, h_{t-1}, x_t] \; + \; b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

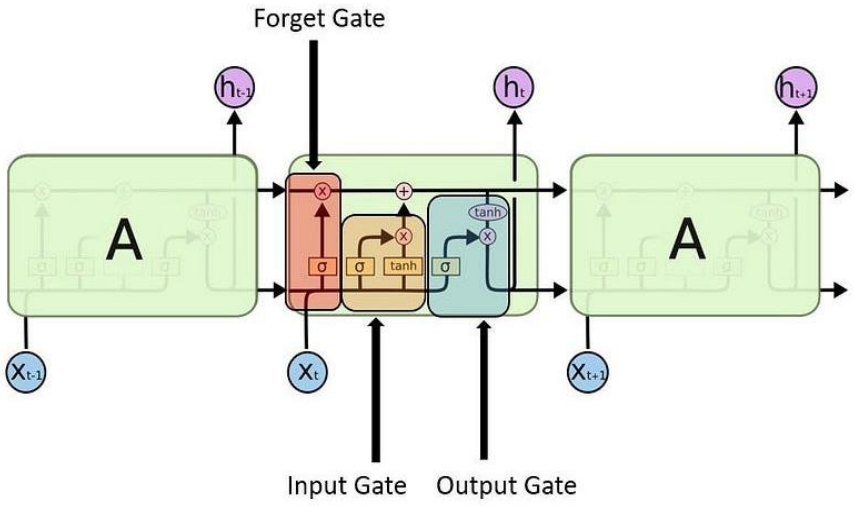


Fig-2.13- LSTM Gates

2.5 Envisioning the Visual World: Computer Vision

2.5.1 Introduction :

Computer vision is one of the most important fields of artificial intelligence (AI) and computer science engineering that makes computer systems capable of extracting meaningful information from visual data like videos and images. Further, it also helps to take appropriate actions and make recommendations based on the extracted information.

2.5.2 Introduction To OpenCV :

Computer vision engineers require in-depth knowledge of machine learning and deep learning concepts. There are so many programming languages that can be used in this domain, but Python is among the most popular. However, one can also choose OpenCV with Python, OpenCV with C++, or MATLAB to learn and implement computer vision applications.

OpenCV with Python could be the most preferred choice for beginners due to its flexibility, simple syntax, and versatility. Various reasons make Python the best programming language for computer vision, which is as follows:

- Easy-to-use
- Most used programming language

2.5.3 How does Computer Vision Work? :

Computer vision is a technique that extracts information from visual data, such as images and videos. Although computer vision works similarly to human eyes with brain work, this is probably one of the biggest open questions for IT professionals: How does the human brain operate and solve visual object recognition?

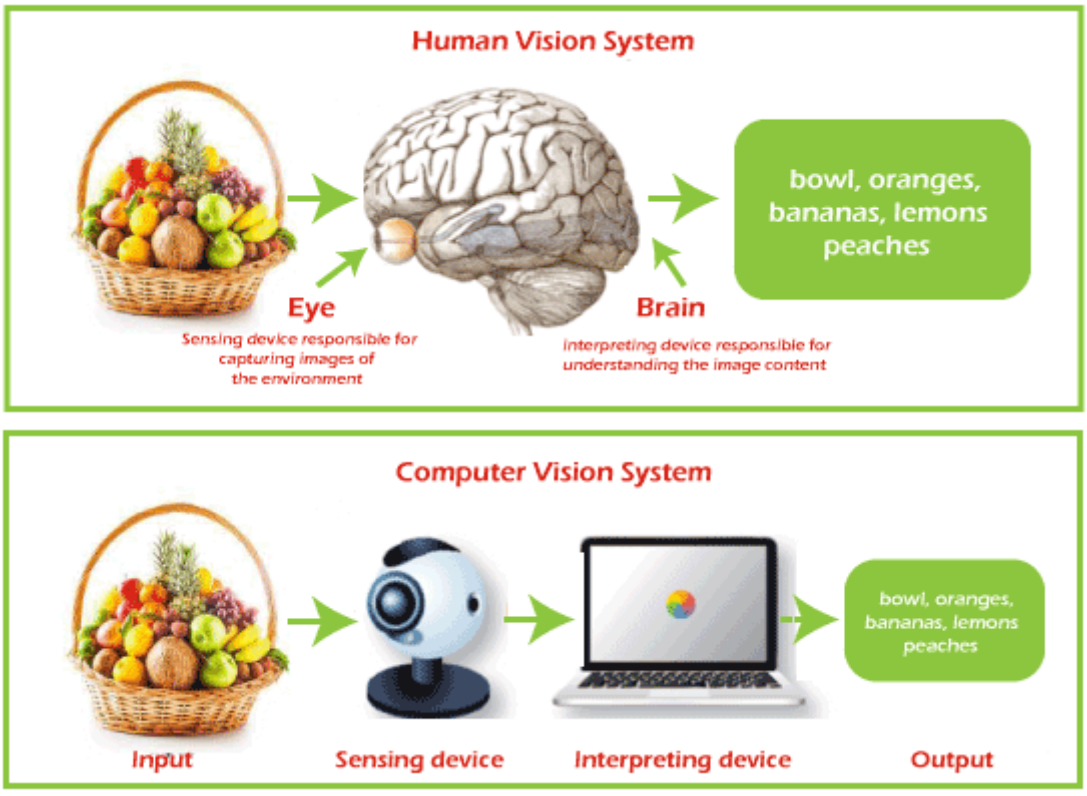


Fig-2.14- *Comparision of Working of a Human Vision system and a Computer Vision System*

On a certain level, computer vision is all about pattern recognition which includes the training process of machine systems for understanding the visual data such as images and videos, etc.

Firstly, a vast amount of visual labeled data is provided to **machines to train it.** This labeled data enables the machine to analyze different patterns in all the data points and can relate to those labels. As a result, **this computer vision model can now accurately detect** whether the image contains a dog or not for each input image.

2.5.4 Applications of computer vision :

- Retail (e.g., automated checkouts): Computer vision is also being implemented in the retail industries to track products, shelves, wages, record product movements into the store, etc. This AI-based computer vision technique automatically charges the customer for the marked products upon checkout from the retail stores.

- Facial recognition: Computer vision has enabled machines to detect face images of people to verify their identity.

- Healthcare and Medicine: Computer vision has played an important role in the healthcare and medicine industry.

2.6 Imagining the Unseen: Generative Adversarial Networks (GAN)

2.6.1 Introduction :

A Generative Adversarial Network or GAN is defined as the technique of generative modeling used to **generate new data sets based on training data sets.** The newly generated data set appears similar to the training data sets.

2.6.2 Components of GAN

Generative adversarial networks are primarily made from two components, i.e., generator and discriminator. As the name suggests, the generator generates a fake output of

unseen data based on training data sets and makes the discriminator fool to understand this fake data as accurate.

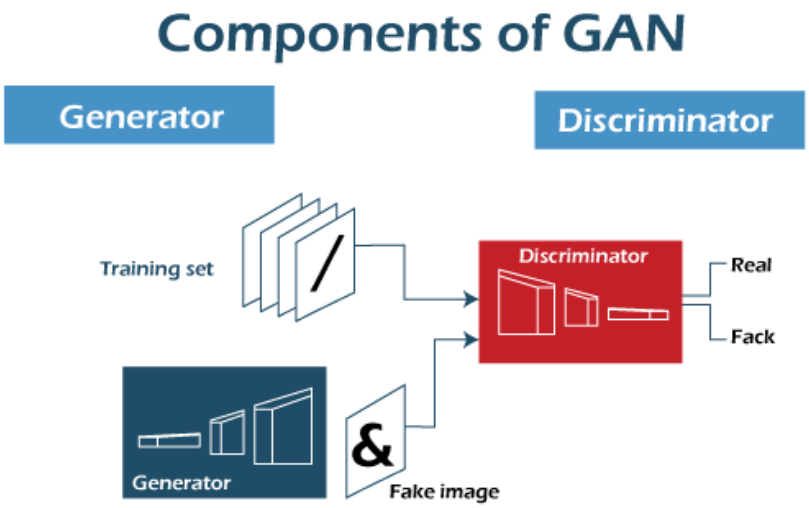


Fig-2.15- *Components Of GAN*

- Discriminator: It is used as a supervised machine learning approach in which a simple classifier is appointed to discriminate between real and fake data. Although, it is trained on actual training data sets and gives feedback to the generator.
- Generator: Unlike the discriminator, the generator is an unsupervised machine learning method used to generate fake samples based on actual training data sets. It is also a neural network with hidden layers, activation, and loss function.

  Further, the generator primarily focuses on generating fake data based on feedback given by the discriminator and makes the discriminator fool so that it cannot identify the difference between actual output and generated output by the generator.

2.6.3 Training of Generative adversarial networks (GANs)

Step-1: Identify the actual problems: This is essential in working on real-time projects. If you can identify the actual problems, you can only solve this efficiently.

Step-2: Choose appropriate GAN architecture: Although there are so many architectures of GANs exist, such as DCGAN, Conditional GAN, Unconditional GAN, Least Square GAN, we have to define which type of GAN architecture we are using in our project.

Step-3: Give training to discriminator on real data sets:

The discriminator is always given training on real data sets, and it only contains a forward path mechanism and does not follow backpropagation in n epochs.

Step-4: Provide training to the generator: The training process of the generator starts with the introduction of some fake inputs. Initially, we give some fake input to the generator, but later it generates some fake output by adding some random noise.

Step-5: Provide training to discriminators on fake inputs: In this step, we pass the samples to discriminators to predict whether the data is real or fake. Further, provide feedback received by decimators to generators again to do modifications in the samples.

2.6.4 Applications of GANs

Below are a few most famous applications of generative adversarial networks (GANs), which are as follows:

- Fashion, art, and advertising
- Science
- Video games
- Audio synthesis
- Transfer learning

CHAPTER 3 – RESULTS AND DISCUSSIONS

3.1 Project Overview and Implementation

3.1.1. Project Scope and Goals

The SHINJITE DRINKS Ordering Virtual Interface project was conceived with the vision of redefining the traditional coffee ordering experience. The scope of the project encompassed the development of an interactive and intuitive interface that leveraged computer vision technology to recognize hand gestures and allow customers to place coffee orders effortlessly. The specific goals included:

- Create a novel drink ordering system that replaces traditional menus with a gesture-based interface.

- Implement real-time hand detection and tracking using computer vision to enable customers to make choices through hand gestures.

- Design an engaging and user-friendly graphical interface to guide customers through the drink selection process.

- Streamline the order placement process to enhance efficiency and reduce errors.

The innovative aspect of the project lay in its application of computer vision for an everyday task like ordering tea or smoothies. By eliminating the need for physical menus and touchscreens, it aimed to provide a unique and contactless experience in line with modern safety standards.

3.1.2 System Architecture and Components

The SHINJITE DRINKS Ordering Virtual Interface system architecture comprised several components that seamlessly integrated to create a cohesive user experience. The key components included:

- **cv2 and opencv-python Libraries :** These formed the foundation for computer vision capabilities, allowing for real-time video processing and hand detection.

- **cvzone Library :** Utilizing the HandDetector module from the cvzone library enabled robust hand detection and tracking.

- **Graphical User Interface (GUI) :** The GUI served as the user's primary interaction point, providing visual cues and feedback.

- **Gesture Recognition Algorithm :** Custom gesture recognition algorithms were developed to map hand configurations to specific coffee preferences.

The implementation of the hand detection and tracking system was a critical aspect of the project's success. By harnessing the cvzone library, the system was capable of detecting both left and right hands, enabling customers to interact naturally without constraints.

3.1.3 Interface Design and Modes

The design of the graphical user interface (GUI) was meticulously crafted to ensure an intuitive and enjoyable user experience. The layout featured clear and concise visual elements that guided customers through the coffee ordering process. The interface was divided into four distinct modes:

- **Drink Selection Mode :** In this mode, customers could choose from three drink options: Lemonade, Bubble Tea, or Chocolate Smoothie. The system recognized the choice based on specific hand gestures.

Fig 3.1 – *Mode 1 and Mode 2 GUI Design*

- **Sugar Level Mode :** Customers progressed to this mode to specify their desired sugar level. The available choices were one spoon, two spoons, or three spoons of sugar.

- **Cup Size Mode :** Next, customers selected their preferred cup size from options such as small, medium, or large.

- **Order Placement Mode :** After making all the selections, customers reached the **Order is placed** screen, confirming their choices and completing the order.

Fig 3.2 - *Mode 3 and Mode 4 GUI Design*

The distinct modes helped streamline the decision-making process, making it user-friendly and approachable. The careful consideration of user interface design contributed significantly to the project's success, as it allowed customers to navigate effortlessly through their coffee preferences.

3.2 Hand Gesture Recognition and Mode Selection

3.2.1 Hand Detection and Tracking

The foundation of the Shinjite Drinks Ordering Virtual Interface's success rested on its ability to accurately detect and track users' hands in real time. The HandDetector module from the cvzone library played a pivotal role in achieving this. This module offered advanced hand detection and tracking capabilities, capable of distinguishing between left and right hands.

Fig 3.3 – *Real-Time Hand Recognition*

3.2.2. Gesture Mapping and Timing

The translation of hand gestures into meaningful selections required a robust mapping system. For instance, a displayed index finger corresponded to selecting a Latte, while showing both

index and middle fingers indicated a preference for Black coffee. Furthermore, the display of all three fingers denoted a choice of Tea. The timing of gesture recognition was strategically set to approximately 4.78 seconds, striking a balance between user comfort and system responsiveness. This duration allowed users sufficient time to make deliberate gestures while preventing undue delays in progressing through the modes.

3.2.3 User Experience and Interaction

The success of the Shunjite Drinks Ordering Virtual Interface hinged on delivering an exceptional user experience. Central to this experience was the responsiveness and accuracy of gesture recognition. Throughout the project's development, rigorous testing and optimization were conducted to ensure that gestures were detected promptly and accurately, even in varying lighting conditions.As a result, the system achieved consistent and reliable gesture recognition, contributing significantly to the overall user satisfaction and usability of the interface.

3.3 Queue Management and Order Display

3.3.1. Real-time Queue Visualization

An essential feature of the Shinjite Drinks Ordering Virtual Interface was its real-time queue visualization. As customers made their selections, their choices were dynamically displayed in a queue at the bottom of the GUI. This queue provided customers with immediate visual feedback on their choices, ensuring transparency and clarity throughout the ordering process.

Fig 3.4 – *Queue Visualisation*

3.3.2 Implementation of "Order Placed" Screen

Upon completing all selections, customers reached the final "Order is placed" screen. This screen served as the ultimate confirmation of their order. It displayed a summary of the chosen coffee preferences, including the selected drink, sugar level, and cup size. The "Order is placed" screen was designed to ensure that customers had a final opportunity to review their choices before confirming their order.

3.4 System Reset and User Interaction

3.4.1. Integration of Reset Mechanism

To further enhance user control and convenience, a system reset mechanism was seamlessly integrated into the Coffee Ordering Virtual Interface. Customers could reset the ordering process at any point by simply showing a “thumb“ gesture to the camera. The thumb gesture served as a universal reset command, enabling customers to return to the initial "Select your drink" mode.

3.4.2 Enhancing User Engagement

The interactive and intuitive nature of the Shinjite Drinks Ordering Virtual Interface contributed significantly to user engagement. By enabling customers to interact with the interface using natural hand gestures, the system fostered a sense of engagement and participation. Customers were actively involved in the coffee selection process, which enhanced their overall experience.

CHAPTER 4 - CONCLUSION AND FUTURE SCOPE

4.1 Conclusion

The Shinjite Drinks Ordering Virtual Interface represents a successful fusion of technology and user experience, redefining the act of ordering drinks. Leveraging computer vision and intuitive hand gestures, this interface has created an engaging and memorable platform for customers. Through meticulous design and seamless integration of components, we have not only streamlined the ordering process but also empowered users with flexibility and control.

The project has demonstrated the accuracy of hand gesture recognition, enhancing user interaction and satisfaction. Real-time queue visualization and visual feedback mechanisms have contributed to a transparent and organized ordering experience. The implementation of a reset feature underscores our commitment to user-friendliness, allowing customers to correct preferences effortlessly.

This innovative interface sets a new standard for coffee ordering systems, promoting user engagement and providing a glimpse into the future of interactive user interfaces. As technology continues to evolve, this project serves as a testament to the boundless potential of human-computer interaction in enhancing everyday experiences.

The project demonstrated the potential of computer vision and user-centric design to enhance everyday tasks and provide innovative solutions to modern challenges.

4.2 Future Scope

4.2.1 Enhanced User Interaction and Personalization

4.2.1.1 Gesture Customization:

- Explore the possibility of allowing users to customize and define their own hand gestures for coffee preferences, providing a highly personalized ordering experience.

4.2.1.2 Voice Integration:

- Integrate voice recognition alongside hand gestures to cater to users with different preferences, making the interface more inclusive and accessible.

4.2.1.3. User Profiles:

- Implement user profiles that store individual preferences and ordering history, enabling a more tailored and efficient ordering process for returning customers.

4.2.2 Expansion Beyond Coffee Ordering

4.2.2.1 Menu Diversification:

- Expand the virtual interface to include a wider range of menu items, such as snacks, desserts, and other beverages, increasing the scope of offerings for customers.

4.2.2.2 Integration with Payment Systems:

- Incorporate secure payment options directly into the interface, allowing users to complete their orders seamlessly without the need for additional transactions.

4.2.2.3 Multi-language Support :

- Develop multi-language support to cater to a diverse customer base, enhancing accessibility for non-English-speaking users.

4.2.2.4 Integration with Mobile Apps:

- Extend the virtual interface's capabilities by integrating it with mobile applications, enabling users to place orders remotely and pick them up at their convenience.

Fig 4.1 - *Building a faster self-checkout system*

These future scope and expansion initiatives aim to further enrich Shinjite Drinks Ordering Virtual Interface, making it even more versatile, user-friendly, and accommodating for a broader range of customers and menu offerings.

REFERENCES

[1] Jose Fumo, “A Gentle Introduction To Neural Network Series”, https://towardsdatascience.com/a-gentle-introduction-to-neural-networks-series-part-1-2b90b87795bc

[2] Pragati Baheti, “Activation Functions in Neural Networks”, https://www.v7labs.com/blog/neural-networks-activation-functions#:~:text=drive%20V7's%20tools.-,What%20is%20a%20Neural%20Network%20Activation%20Function%3F,prediction%20using%20simpler%20mathematical%20operations.

[3] James Liang, “An Introduction To Deep Learning”, https://towardsdatascience.com/an-introduction-to-deep-learning-af63448c122c

[4] Srivignesh Rajan, “An Introduction to Artificial Neural Networks”, https://towardsdatascience.com/an-introduction-to-artificial-neural-networks-5d2e108ff2c3

[5] Christopher Thomas, “An Introduction To Convolutional Neural Networks”, https://towardsdatascience.com/an-introduction-to-convolutional-neural-networks-eb0b60b58fd7

[6] Kousai Smeda, “Understand The Architecture of CNN”, https://towardsdatascience.com/understand-the-architecture-of-cnn-90a25e244c7

[7] Ronak Vijay, “A Gentle Introduction To Natural Language Processing”, https://towardsdatascience.com/a-gentle-introduction-to-natural-language-processing-e716ed3c0863

[8] Jonte Dancker, “A Brief Introduction To Recurrent Neural Networks”, https://towardsdatascience.com/a-brief-introduction-to-recurrent-neural-networks-638f64a61ff4

[9] Rian Dolphin, “ LSTM Networks | A Detailed Explanation”, https://towardsdatascience.com/lstm-networks-a-detailed-explanation-8fae6aefc7f9

[10] Llija Mihajlovic, “ Everything You Ever Wanted To Know About Computer Vision”, https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e