Artificial Intelligence

AI is not about replacing humans but about augmenting human capabilities

Prepared by
Prasannahariveeresh
S1 - EAC

Table of contents

About Me!

Self-Driven and ambitious student passionate about Artificial Intelligence, Machine Learning, and Computer Vision, seeking proper guidance and platform to explore further and contribute to these fields.

Few Achievements:

- IBM-SkillsBuild, Certified AI developer
- ISRO Design Competition, First prize winner
- Freelance developer for a few startups during my 11th and 12th
- National Level CBSE Science Expo, Participant
- Tamil Nadu Science Federation Science Expo, Winner

Contact Me:

Phone: +91 8925990097
prassijr@gmail.com
jrprasanna.com
github.com/Prasannahariveeresh

Contents of this presentation

- This presentation provides concise explanations of important AI and machine learning concepts, expressing my comprehension with in-depth analysis and insight
- This presentation focuses in explaining what is a model and what is the AI's operating principle.
- Also explains what is a Classifier, gives definitions and explanations for a few classifiers that I am familiar with.
- Explaining and Comparing Supervised Learning and Unsupervised Learning
- Explaining the working principles of ChatGPT and Gemini

“Artificial Intelligence is the new electricity. Just as electricity transformed industry after industry a hundred years ago, AI will now do the same.”

—Andrew Ng

Co-founder of Coursera and Adjunct Professor at Stanford University

01

Working Principle of AI

A brief about what is AI, how it works and what is a Model

Principle of AI

01 Understanding Dataset

02 Data Preprocessing

03 Selecting Model

04 Training Model

05 Prediction or Classification

Understanding the DataSet

- In this step we gather all the data which can be image, audio, Database, CSV or any other format
- The sources of these data can be from sensors, internet, databases or manual inputs
- The problem statement should be understood and based on that classification or prediction can be done
- The quantity and the quality of the dataset is crucial to feed in a lot of data to the model

Data Preprocessing

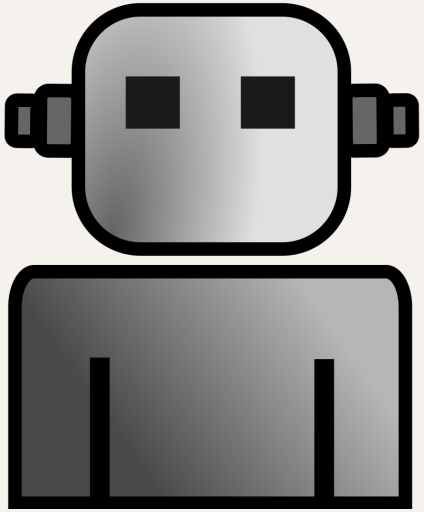
- Usually raw data might not contain certain values or with some sort of error
- In this step, we clean the data, by dropping duplicates, missing values and etc.
- We can go with data augmentation in case of very less data is given, this step can be done to improve
  the accuracy of the model
- Splitting the dataset into training (70-80%) and testing (30-20%) is crucial

Selecting the Model

- From the understanding the data, we will come to know that the given dataset is a classification, prediction, regression or clustering
- So a suitable algorithm should be selected. Ex. Support Vector Machines, Decision Trees, Neural Networks, etc.
- Based on their training, metrics will be evaluated and prior model can be selected

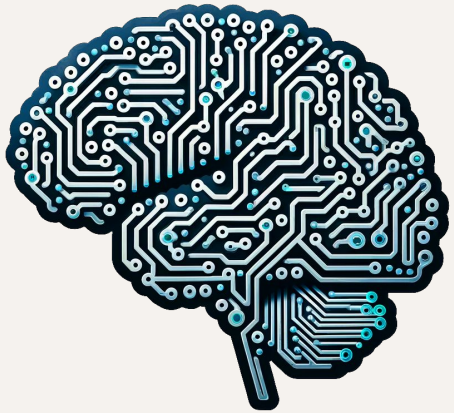
Training the Model

- Training is the process where we will fit our cleaned and preprocessed data to our selected model
- Models can also be selected based on their performance metrics like
  - Accuracy, Precision, F1-Score: For classification datasets
  - R-Squared, MAE, MSE: For regression datasets



Predicting Or Classifying

- After training the model will be evaluated with the metrics and the best model will be selected. The model will be evaluated with the testing data
- Now the AI model is capable of classifying or predicting on new data
- After furthermore testing it can be deployed in a server and letting the users access it
- This is the final stage of solving an AI problem



02

Classifiers

A brief on Classifiers, with examples

What is a Classifier ?

- **Classifier** is an algorithm that categorizes or classify the data into given class labels or categories
- They classify data to given labels based on their input features
- Classifiers are of three types
  - **Binary Classifiers -** SVMs, Binary decision tree, etc.
  - **Multi label Classifiers -** Neural Networks, k-Nearest Neighbours, etc.
  - **Ensemble Classifiers -** RandomForest (collection of decision trees), etc.

Visual representation of Classifier

Input

Classifier

Car ✅

Accuracy: 86.663

Bike

Accuracy: 2.917

Boat

Accuracy: 10.42

Examples of Classification Algorithms

- **Logistic Regression:** Used for binary classification
- **Decision Trees:** Used for classification as well as regression
- **RandomForest:** Used for classification as well as regression
- **k-Nearest Neighbours (kNN):** Mostly used for classification, but can be used for regression too
- **Support Vector Machines (SVMs):** Used for classification as well as regression
- **Neural Networks**

And many more...

Usage of Classifiers in Real Life

- **Spam Detection:** Gmail often classifies the mail as spam, by using classifiers
- **Sentimental Analysis:** Social media comments are classified into positive, negative and neutral based on user's input
- **Image Detection:** YOLO v8 is a powerful object detection model that detects various objects, animal detection, lane detection, etc are possible
- **Medical Diagnosis:** Detection of brain tumour, cancer and etc. through DICOM files
- **Handwriting detection:** OCR can be built from input received in camera

03

Supervised vs Unsupervised Learning

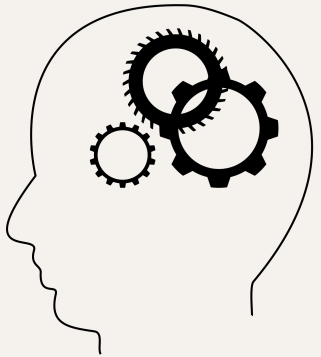
A brief on how do machines learn

What is Supervised Learning ?

- **Supervised learning** is a type of Machine Learning Algorithm where the model is trained on labeled dataset
- The model receives input i.e. the data used to make predictions called the **Input Features (X)** and the output that is associated to input is called **Output Labels (Y)**
- During training, the model makes prediction based on X and compares the predictions with the actual labels. The difference between prediction and labels is called **Error** and this is used to adjust the model parameter by tweaking learning rate, preventing overfitting or underfitting, etc.

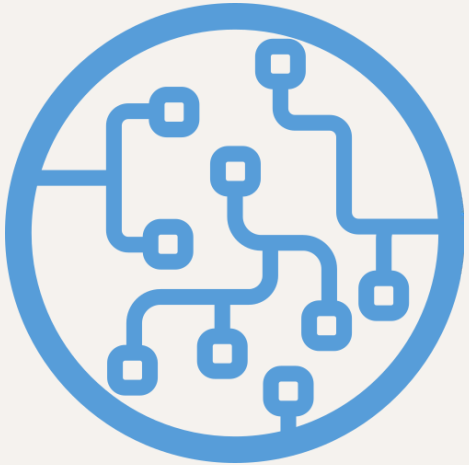
Types of Supervised Learning

- **Supervised Learning** are of two types
  - Regression
  - Classification
- **In Regression,** the output is a continuous value i.e. it is not specified within the given
- **In Classification,** the output is within the given label



What is Unsupervised Learning ?

- **Unsupervised learning** is a type of Machine Learning Algorithm where the model is trained on a non-labeled dataset
- The goal here is to analyse and find the patterns within the input data
- Widely used for **clustering** (eg, customer segmentation)
- This learning method is suited for **Exploratory Data Analysis (EDA),** where grouping can be done without predefined variables



04

How ChatGPT and Gemini works

A brief on what is LLM

What is LLM ?

- **Large Language Model** called as LLM, this model is trained on large amounts of text data to generate human text
- Parameters are the values that the model learns during the training, these LLMs are fed in billions of parameters
- LLMs are trained in such a way that model learns to predict the next word in a sentence
- They also capture grammar, learn important facts and etc.
- OpenAI’s ChatGPT, Google’s Gemini, Meta’s Llama 3 are popular LLMs

Principle of ChatGPT and Gemini

- As discussed earlier, ChatGPT and Gemini are LLM.
- ChatGPT uses GPT-3 and GPT-4 as LLM. GPT (Generative Pre-Trained Transformer) and Gemini uses deep learning approach to learn and produce output
- GPT-3 uses 175 billion parameters whereas Google hasn’t disclosed information about the number of parameters used in Gemini
- They are trained in vast datasets, but Gemini inputs images also

Thank you !

Do you have any questions?

prassijr@gmail.com
+91 89259 90097
jrprasanna.com
github.com/Prasannahariveeresh