Machine Learning – BCSE209L
Introduction to Machine Learning

DR. BHARGAVI R

SCOPE

VIT CHENNAI

Course Prerequisites

You are required to have a good understanding of the following

- Linear Algebra
- Calculus
- Probability and Statistics
- Python Programming (NumPy, Pandas, Matplotlib, Scikit-Learn)

Course Components and Grading

Component	Credits	Grading
Theory	3	Assignments - 30 CAT1 - 15 CAT2 - 15 FAT - 40
Lab	1	CAM - 60 FAT - 40

Course Goals

Equip you with concepts of ML, tools and techniques to design and implement Machine learning models for real time applications.

1. Understanding fundamentals, algorithms, analysis.
2. Hands on - Implementation (in Python) and analysis with toy datasets
   - Platform – Kaggle
   - Google Colab

Course page:

https://bhargaviren.github.io/

Course Content – Key Topics

Models

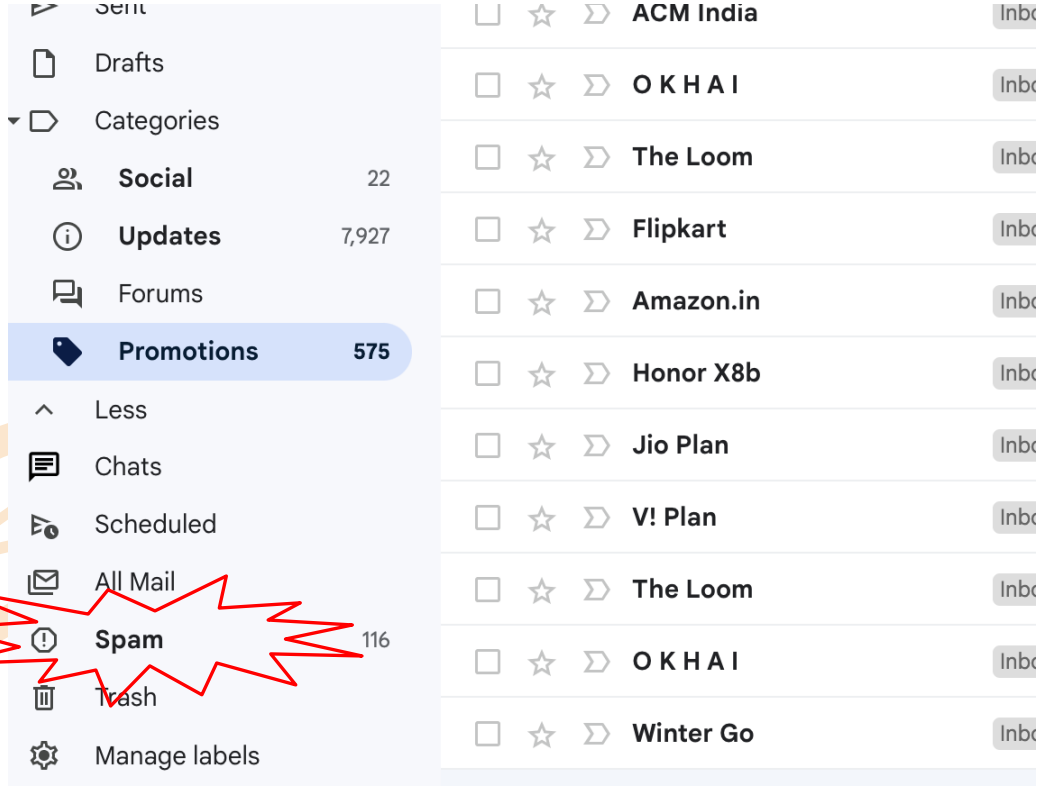
K- Nearest Neighbors Linear Regression Logistic Regression Decision Trees Support Vector Machines (SVM) Neural Networks
K-Means(++), K-Modes Clustering GMM PCA
Probability Based Models: Naïve Bayes
Reinforcement Learning

Concepts

PAC learning Gradient Descent Bias-Variance tradeoff, Overfitting Evaluation Metrics Model Selection Ensemble – Bagging, Boosting

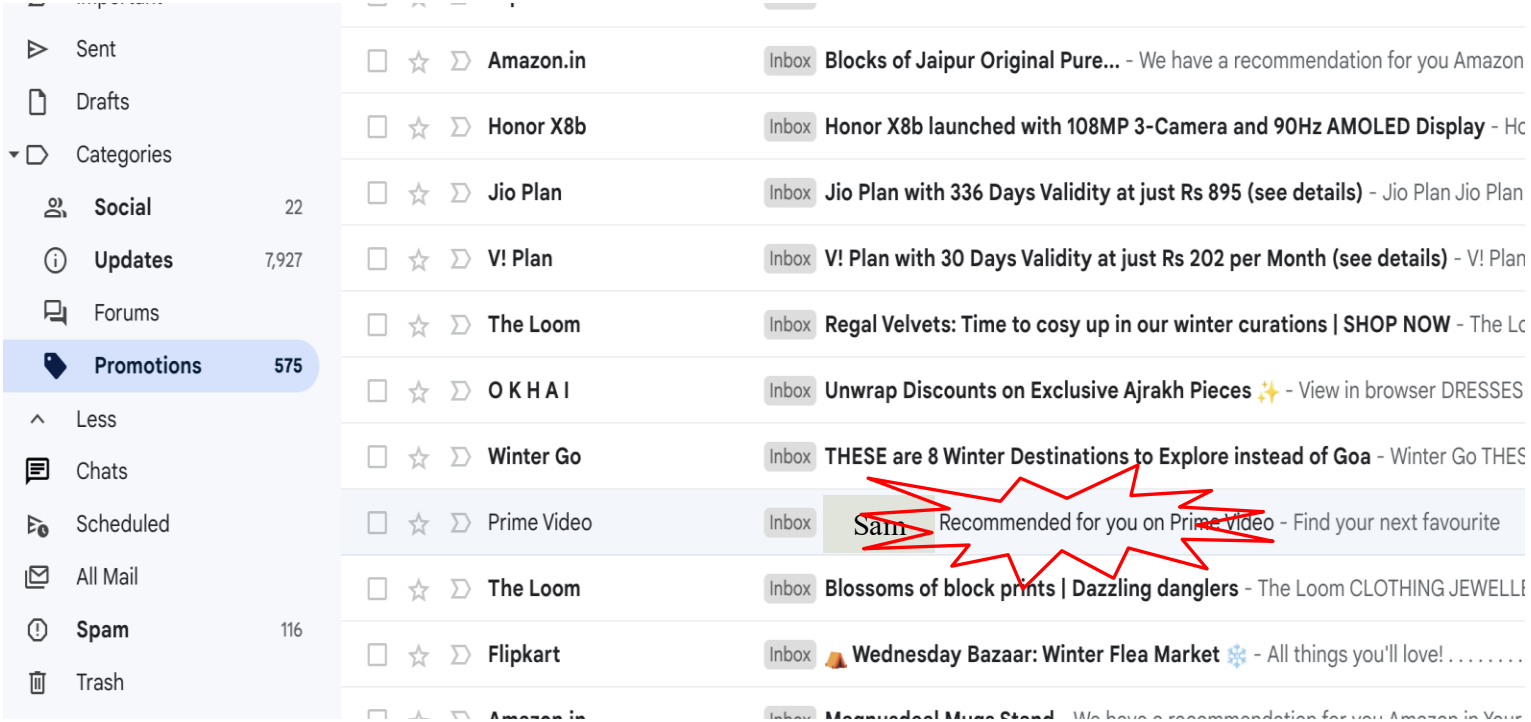
Machine Learning in Daily Life

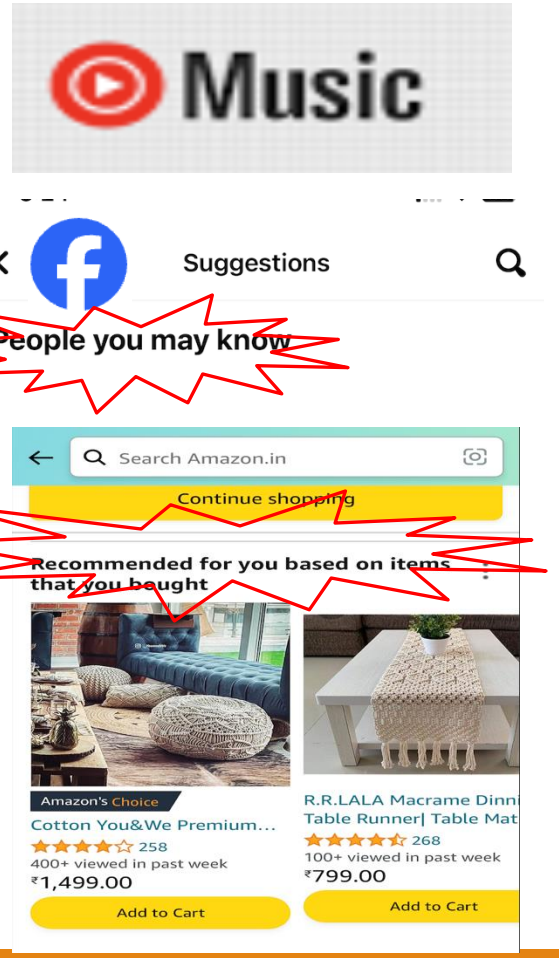
SPAM mail prediction



Machine Learning in Daily Life (cont…)

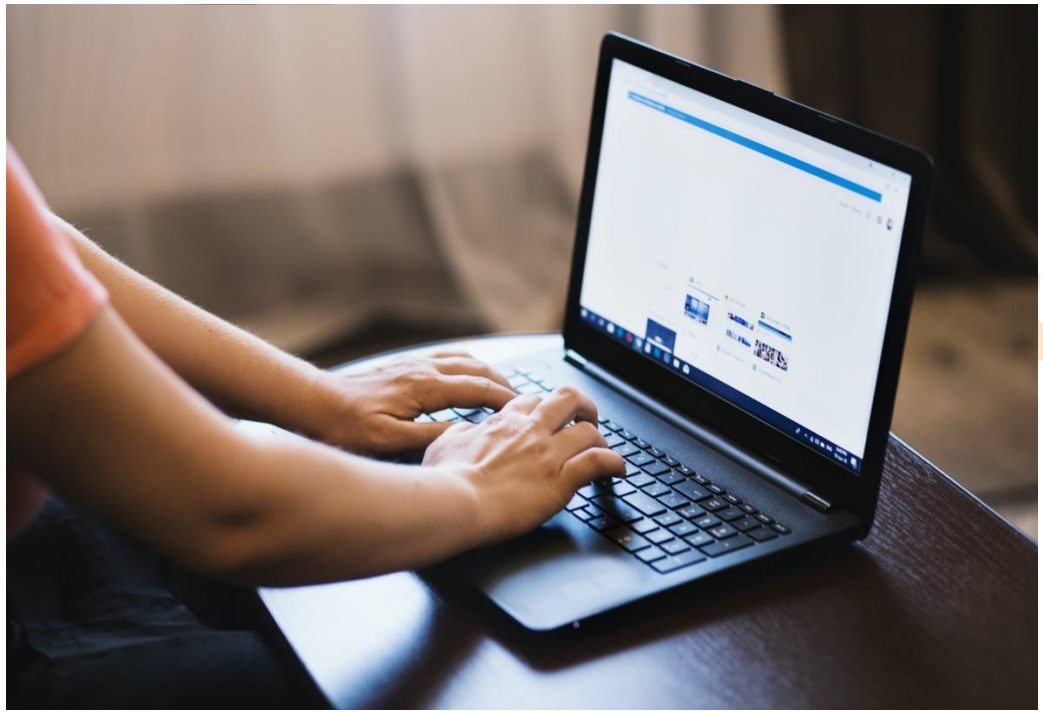
Recommendation systems (Movie/Song/Product/Friends)

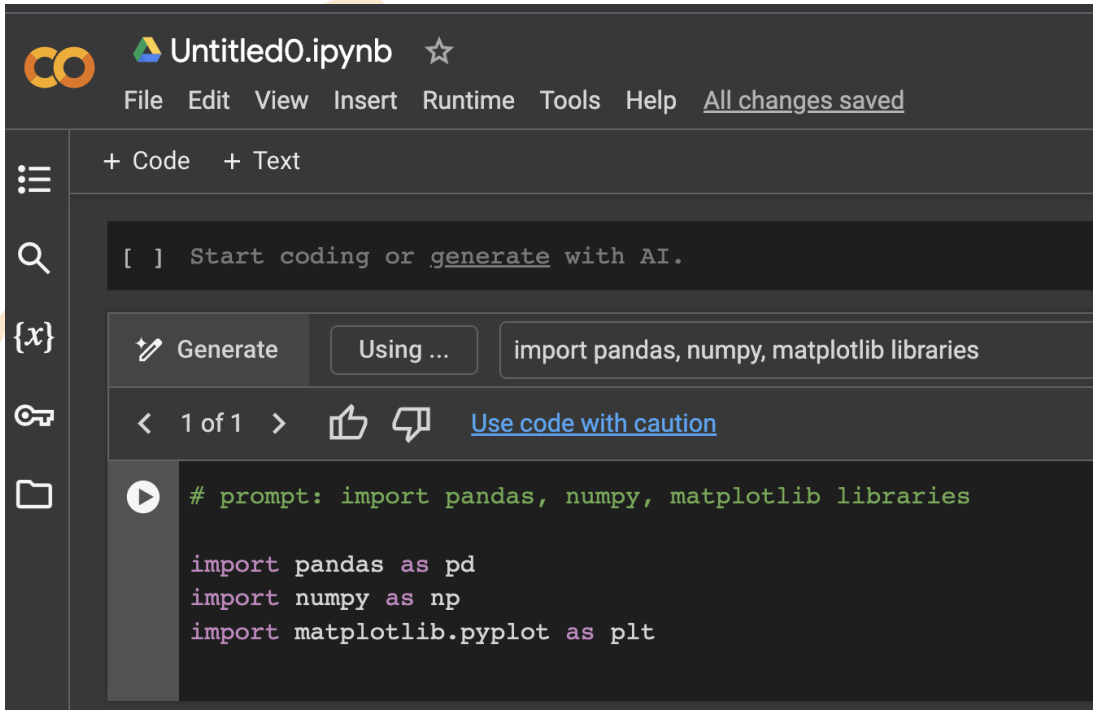




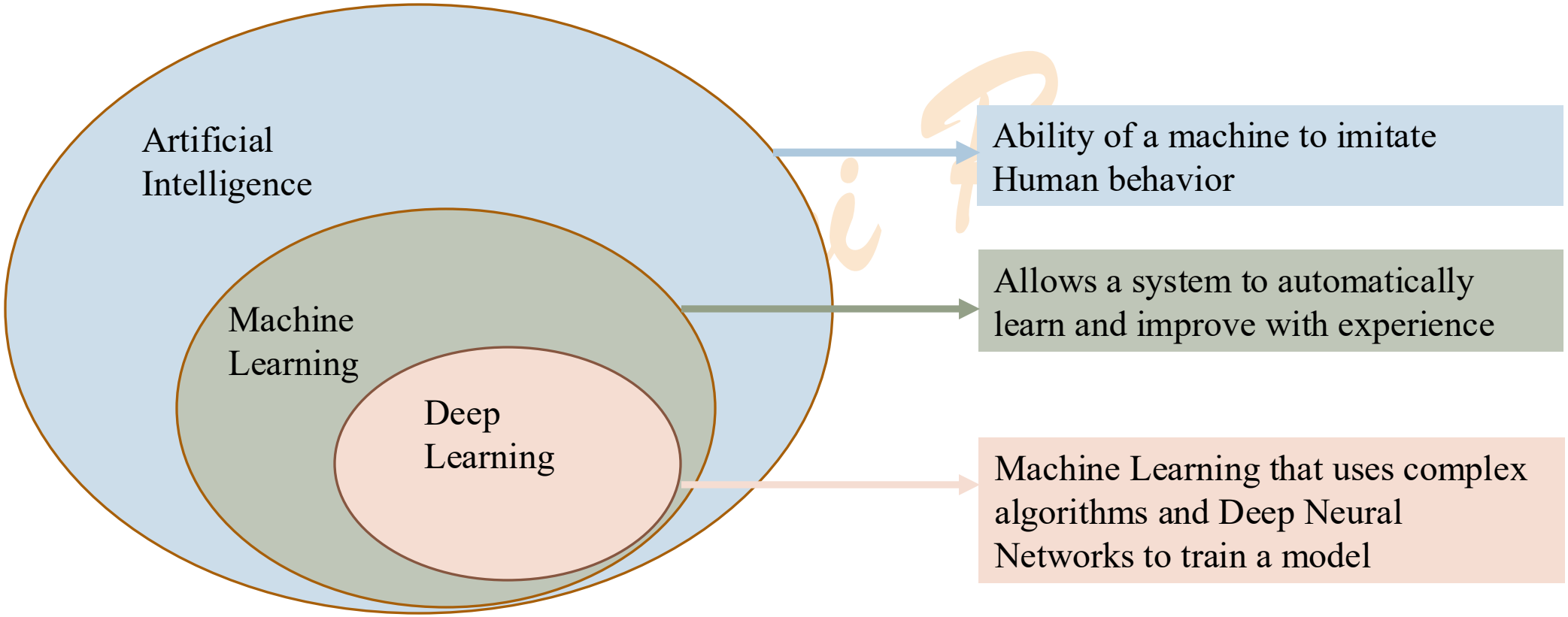
Machine Learning in Daily Life (cont…)

Code generation using AI





AI, Machine Learning and Deep Learning



What is Machine Learning?

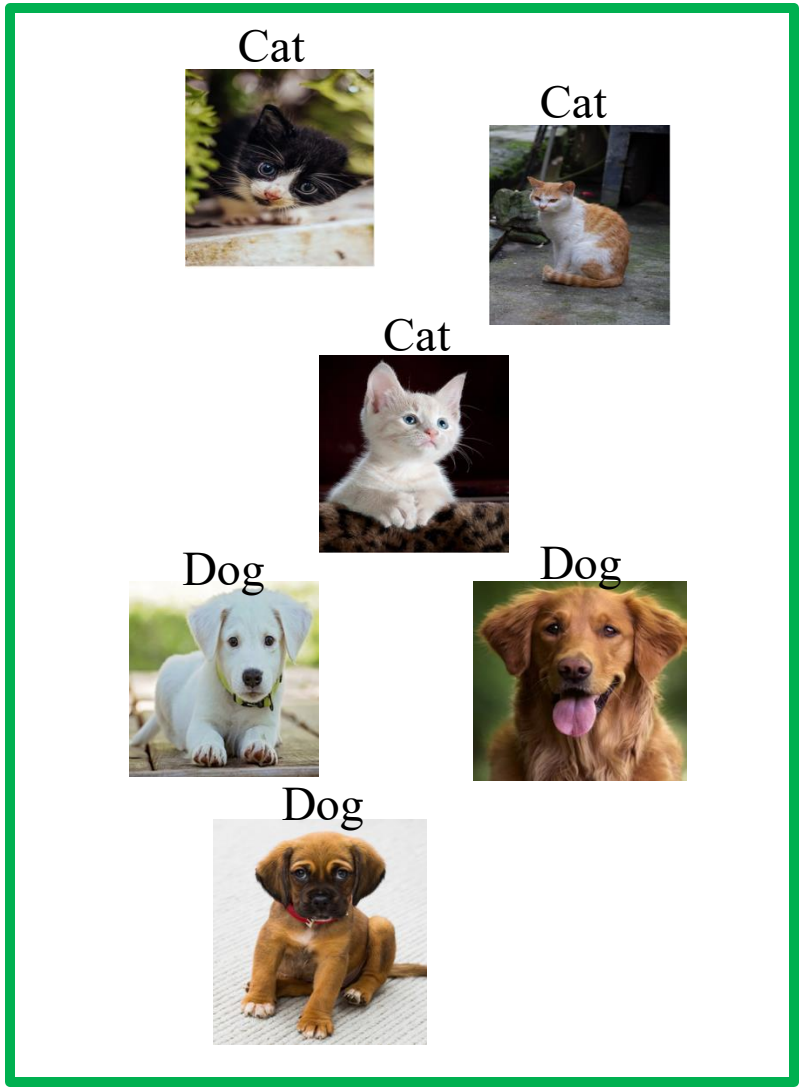
Can you write a program to identify this image?

Can you use the same program to identify this cat?

Then this?

What is Machine Learning? (cont…)

Training Data Set (labeled Examples)



Machine Learning (cont…)

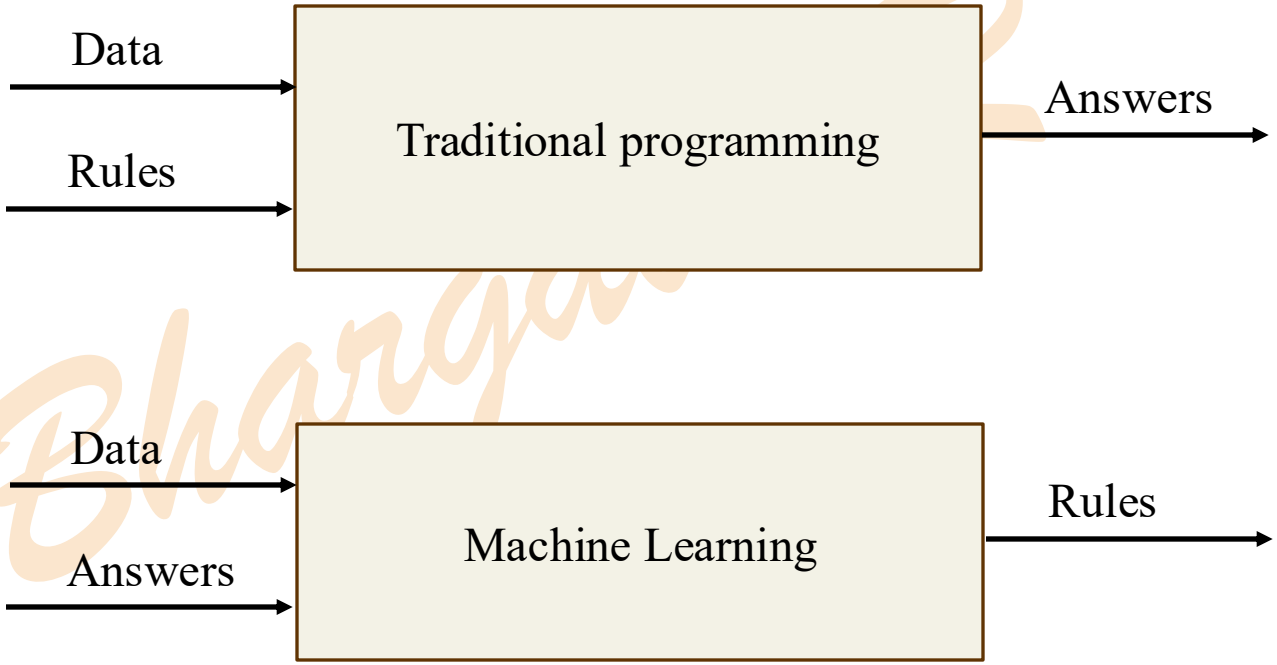
Machine Learning

- The field of study that gives computers the ability to learn ***without*** being ***explicitly programmed***. ~**Arthur Samuel**

- A computer program is said to learn from ***experience*** E with respect to some class of ***tasks*** T and ***performance measure*** P, if its performance at tasks in T, as measured by P, improves with experience E. ~**Tom Mitchell**

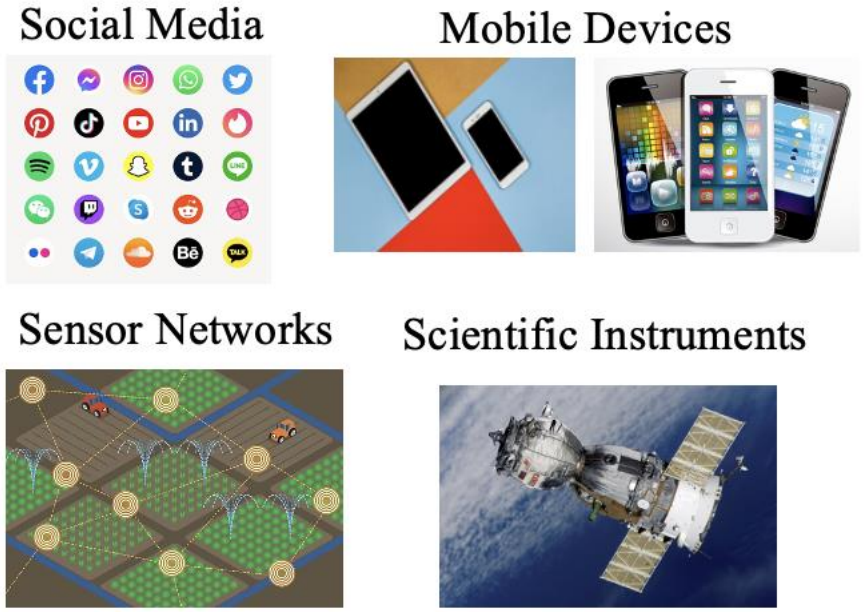
Traditional Programming vs Machine Learning

Traditional programming Vs Machine Learning



Why now?

- Exponential growth of data in recent years
- Who is generating ?
- Is there some valuable information hidden in the data?
- How can we get the insights of this hidden information?
- ***Machine Learning***



Machine Learning - Essentials

Existence of a pattern

Pattern can not be pin down mathematically (hence called Learning from data)

There is data (experience)

Example

Scenario: ABC Supermarket targets to improve its profits

Objective: Improve profits by improving the sales

How ?

Predict

- Which customer is likely to buy which product?
- Which products/brands are preferred by which customers?

Example (cont…)

What is available with ABC?

Transaction Logs

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Invoice ID	Branch	City	Customer	Gender	Product line	Unit price	Quantity	Tax 5%	Total	Date	Time	Payment	cogs	gross mar	gross inco	Rating
2	750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7	26.1415	548.9715	01-05-2019	13:08	Ewallet	522.83	4.761905	26.1415	9.1
3	226-31-3081	C	Naypyitaw	Normal	Female	Electronic accessories	15.28	5	3.82	80.22	03-08-2019	10:29	Cash	76.4	4.761905	3.82	9.6
4	631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	03-03-2019	13:23	Credit car	324.31	4.761905	16.2155	7.4
5	123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8	23.288	489.048	1/27/2019	20:33	Ewallet	465.76	4.761905	23.288	8.4
6	373-73-7910	A	Yangon	Normal	Male	Sports and travel	86.31	7	30.2085	634.3785	02-08-2019	10:37	Ewallet	604.17	4.761905	30.2085	5.3
7	699-14-3026	C	Naypyitaw	Normal	Male	Electronic accessories	85.39	7	29.8865	627.6165	3/25/2019	18:30	Ewallet	597.73	4.761905	29.8865	4.1
8	355-53-5943	A	Yangon	Member	Female	Electronic accessories	68.84	6	20.652	433.692	2/25/2019	14:36	Ewallet	413.04	4.761905	20.652	5.8
9	315-22-5665	C	Naypyitaw	Normal	Female	Home and lifestyle	73.56	10	36.78	772.38	2/24/2019	11:38	Ewallet	735.6	4.761905	36.78	8

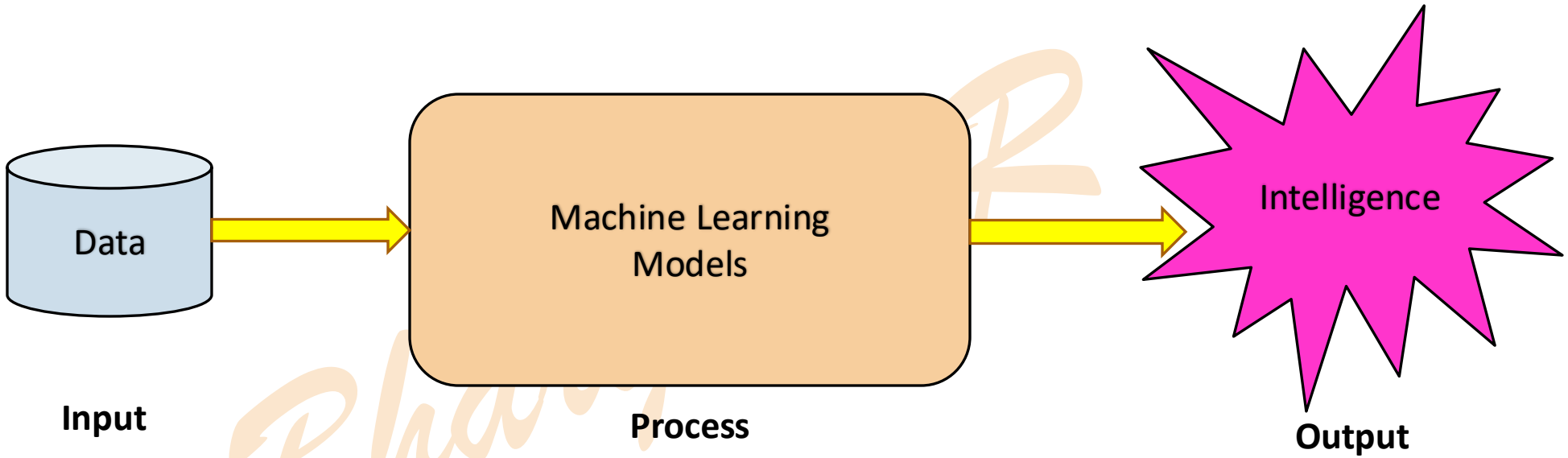
Data Source: https://www.kaggle.com/aungpyaeap/supermarket-sales

What can be done with this transaction logs?

Example - What can be learned from logs?

- Identify the ***items purchased together*** for strategic ***promotional activities***
- ***Segment*** the ***customers*** based on their ***preferences*** and ***target promotional activities***
- ***Predict*** the ***potential buyers*** of specific products using ***demographic information, customer preferences*** etc.,

Machine Learning – Bird's-Eye View



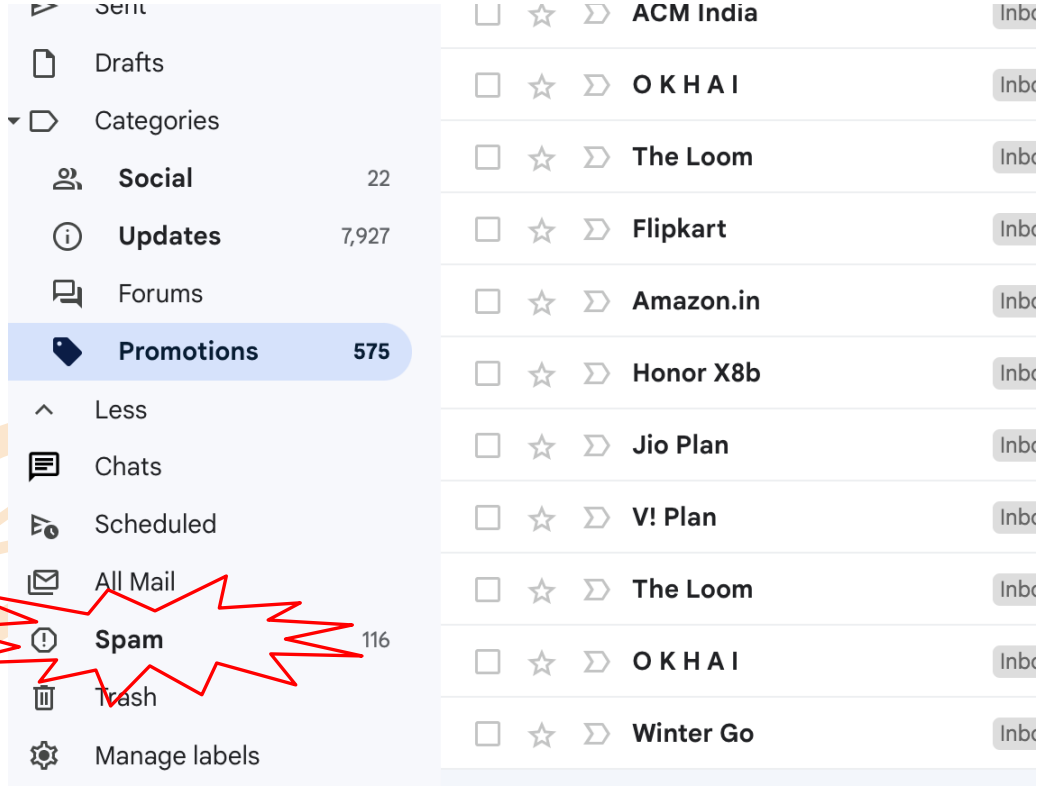
Machine Learning Applications

Machine learning is commonly used in the following applications

- Predictions
- Data Compression
- Generating synthetic data
- Learning series of actions

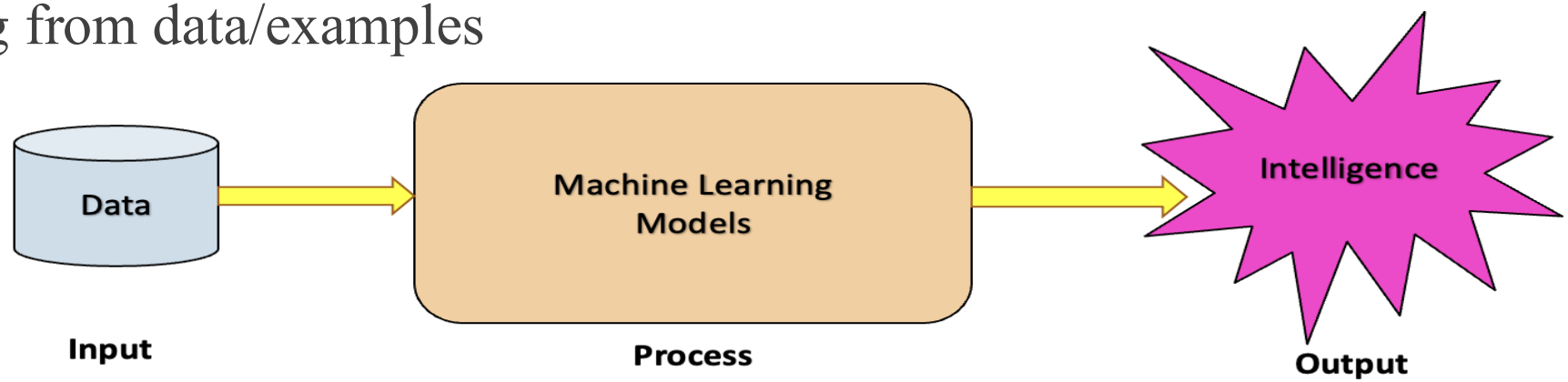
Predictions

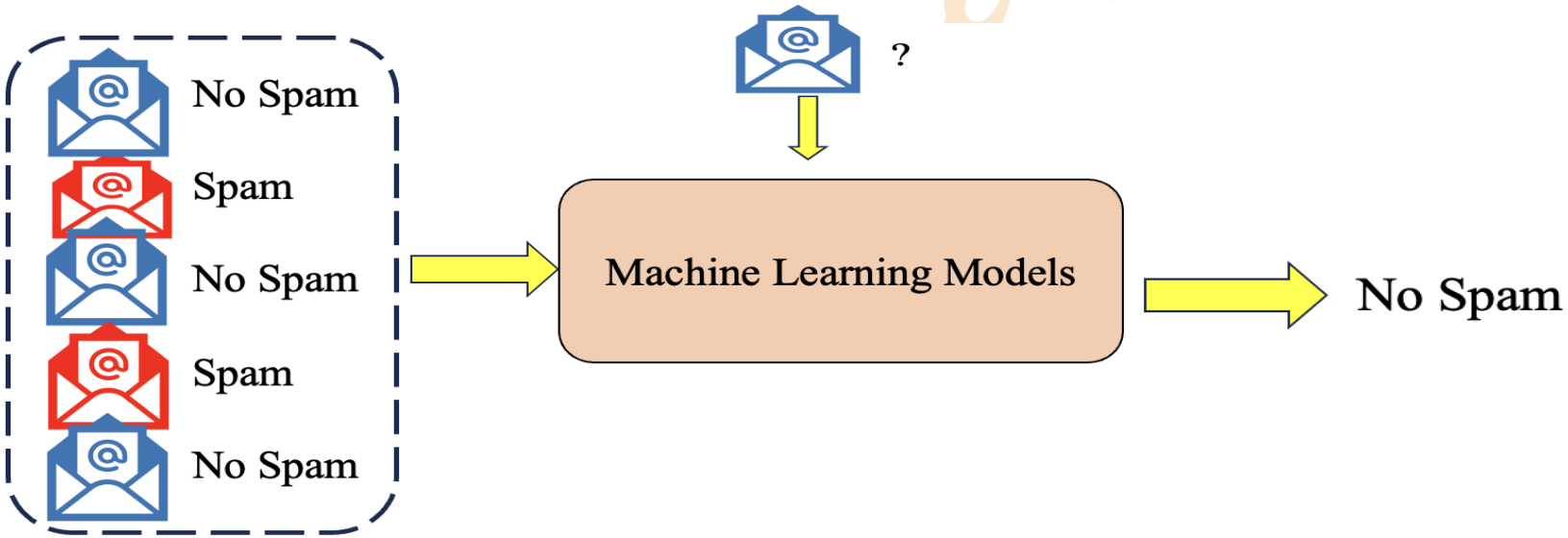
SPAM mail prediction



Predictions (cont…)

Learning from data/examples





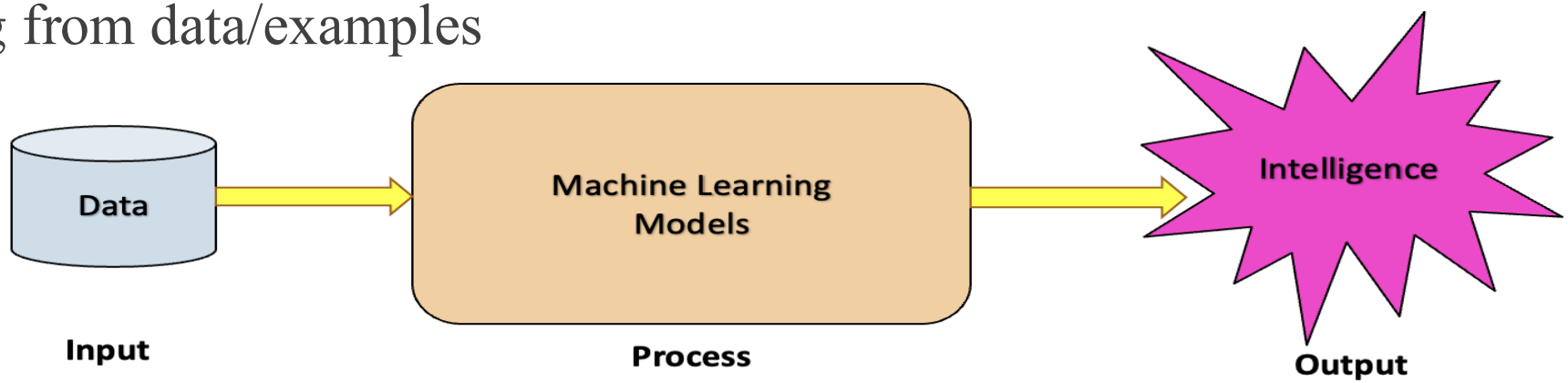
Predictions (cont…)

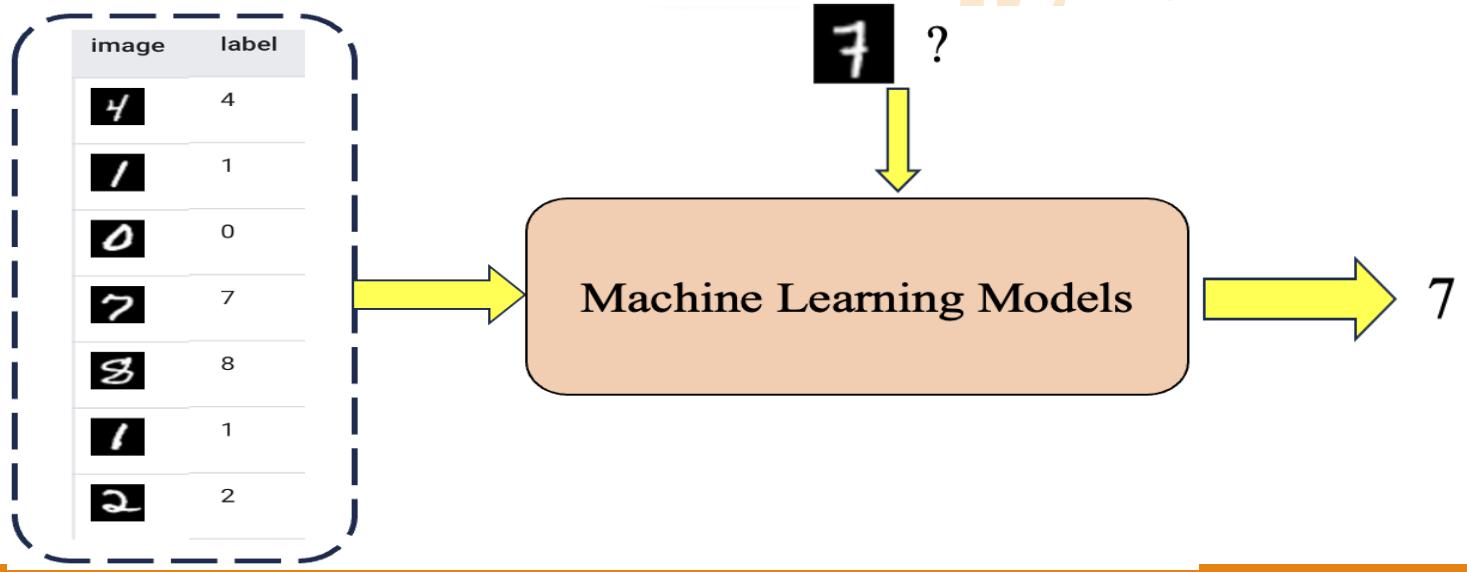
Optical Character Recognition (OCR)

House Number Identification - Google Street View (SVHN)

Predictions (cont…)

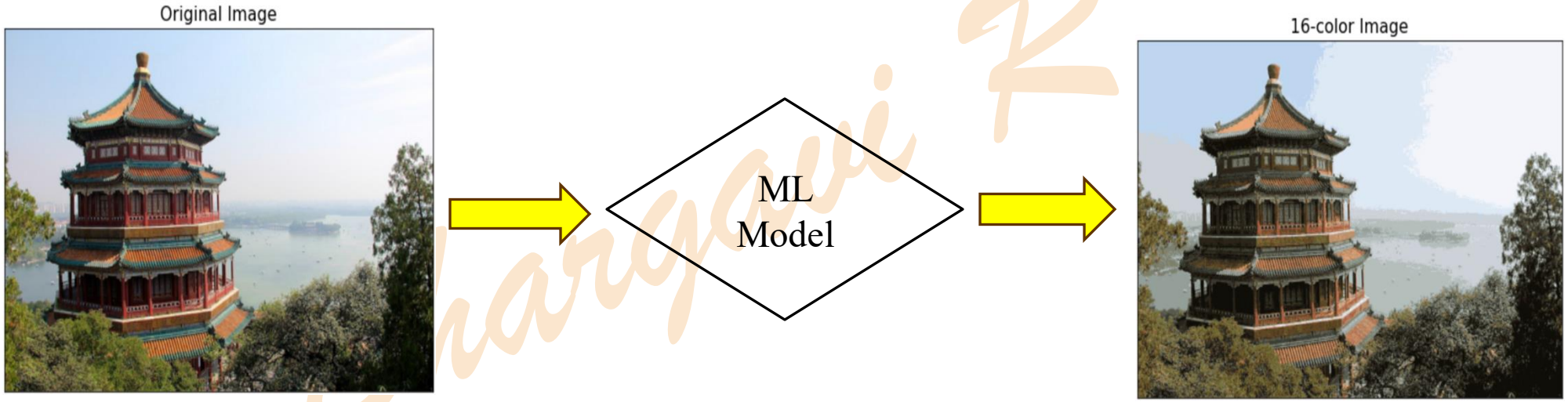
Learning from data/examples





Data Compression

Image compression

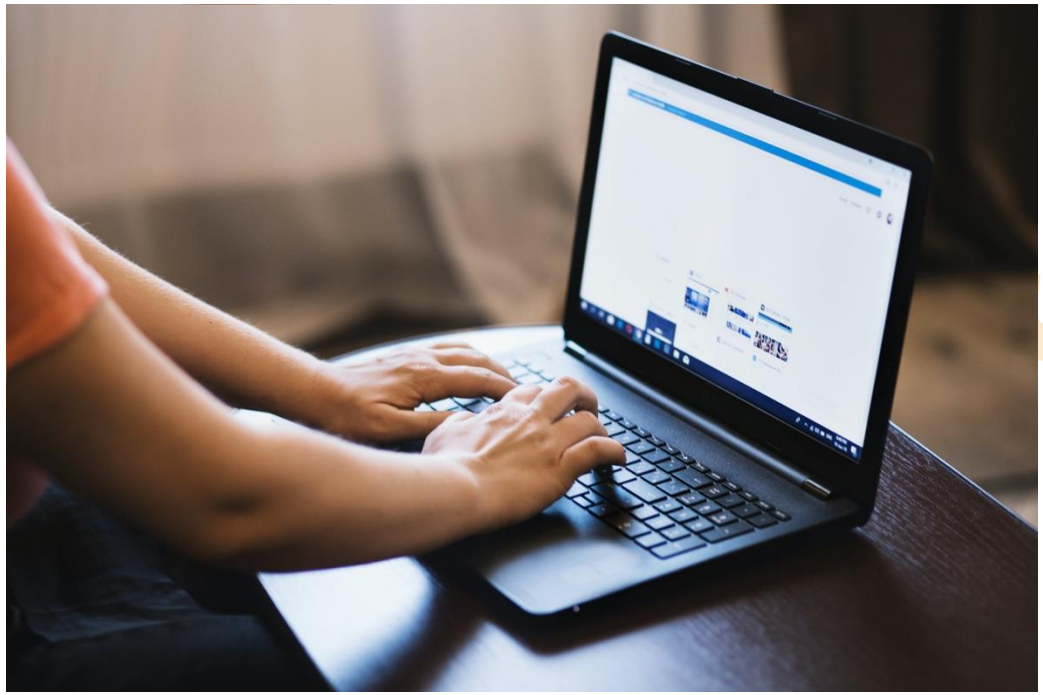


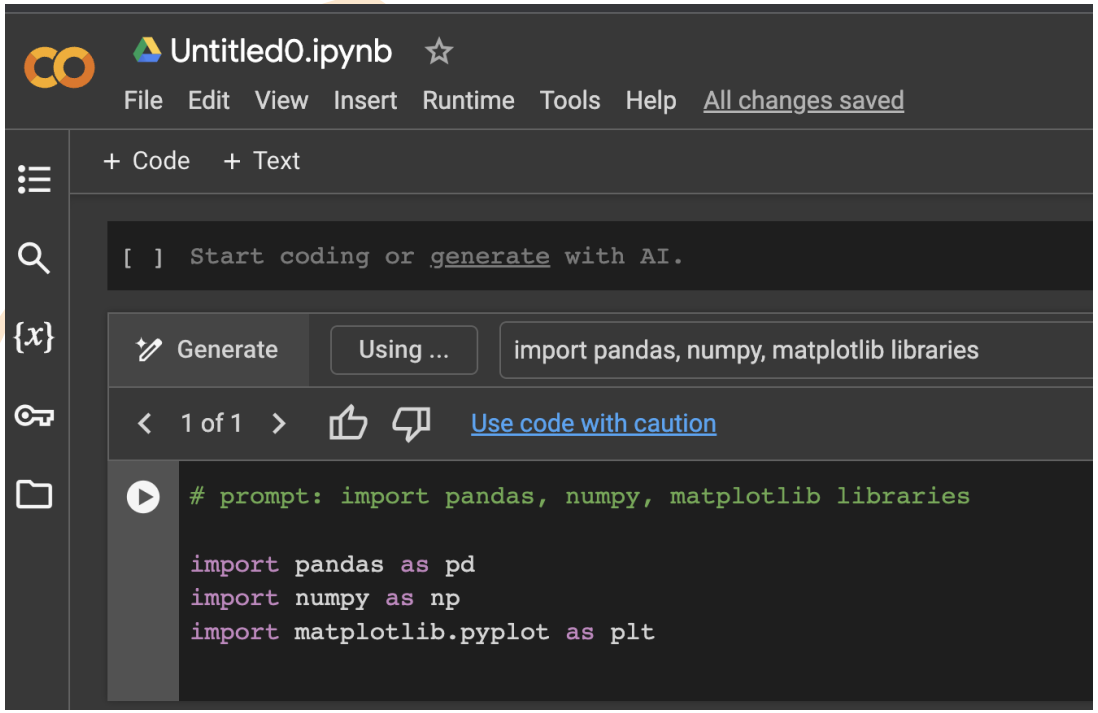
Data Generation

Source: https://www.kdnuggets.com/2020/03/generate-realistic-human-face-using-gan.html

Data Generation (cont …)

Code generation using AI





Learning a series of actions

Computer playing chess

Robo stocking shelves

Machine Learning – Formal Definition

The field of study that gives computers the ability to learn without being explicitly programmed.

~ Arthur Samuel

A computer program is said to learn from ***experience E*** with respect to some class of ***tasks T*** and ***performance measure P***, if its performance at tasks in **T**, as measured by **P**, improves with experience **E**.

~ Tom Mitchell

Task, Performance and Experience

Learning	Task (T)	Performance Measure (P)	Experience (E)
Hand writing recognition	Recognizing and classifying the hand written words	Percent of words correctly classified	Database of handwritten words with correct classification
Dyslexia prediction	Predicting whether a child is dyslexic or not	Percentage of correct predictions	Database of eye movements of children while reading
Robot car driver	Driving a car on public road using vision sensors	Average distance traveled before an error	A sequence of images and steering commands recorded while observing a human driver
Playing Checkers	Playing Checkers game	Percent of games won against the opponent	Playing number of games

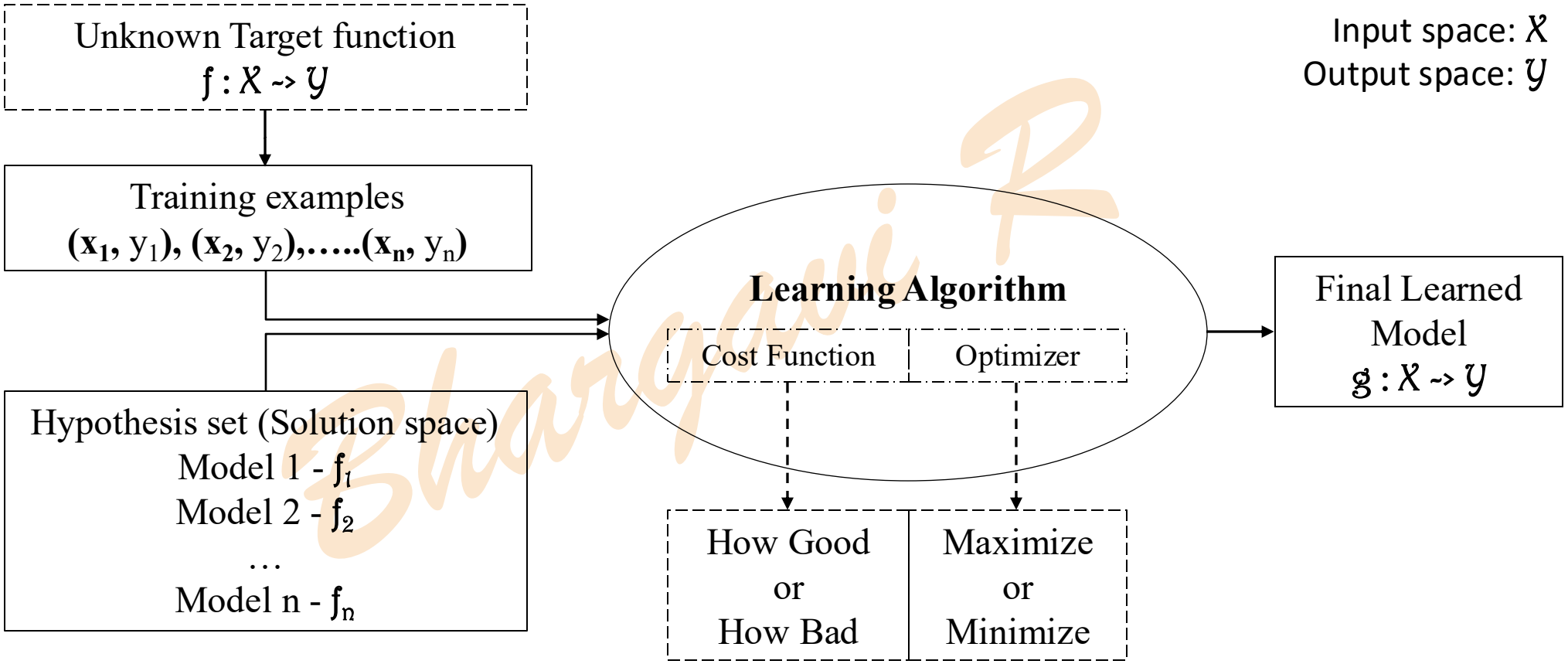
Notations

Symbol	Data type	Description
x	scalar	Ex: 5
x	vector	*An input record* with attributes or features. Ex: Sales info like Product Name, Price, Invoice ID etc., [Apple, 50, 123]
X	matrix	Input sample space. Ex: *Set of all input records* of the sales.
n	scalar	Number of records or instances or observations.
m	scalar	Number of features (dimensions). Ex: If **x** = Product Name, Price, Invoice ID, then *m=3*
K	scalar	Number of outputs (Classes)
$\mathbf{x}_i$	vector	Vector containing i^{th} record. Ex: 5^{th} record

Notations (cont…)

Symbol	Data type	Description
y	vector	Output space. Ex: Ratings of all the products
x_{ij}	scalar	Value at j^{th} attribute in i^{th} record. Ex: Value at 5^{th} column in 2^{nd} record Ex: Value at Price column in 2^{nd} Transaction
f(x)	function	**f : X -> y**. f is a function mapping input space **X** to output space **y** i.e. **y** = f(**x**)
h	hypothesis	A candidate formula or function that maps input to output
H	hypotheses	Set of hypothesis.

Machine learning - Process



Machine Learning - Paradigms

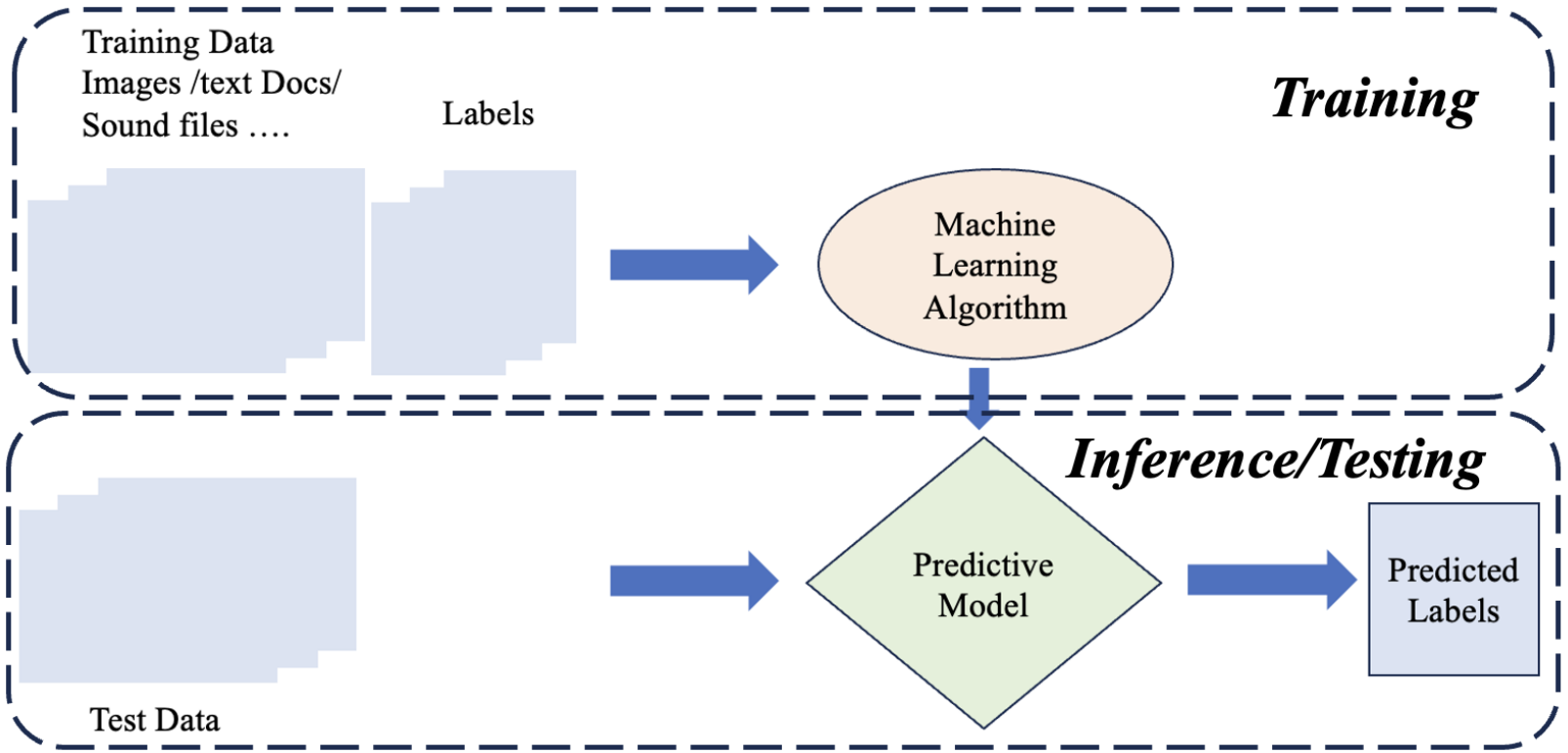
Supervised (Predictive) Learning

Unsupervised (Descriptive) Learning

Reinforcement Learning

Supervised Learning

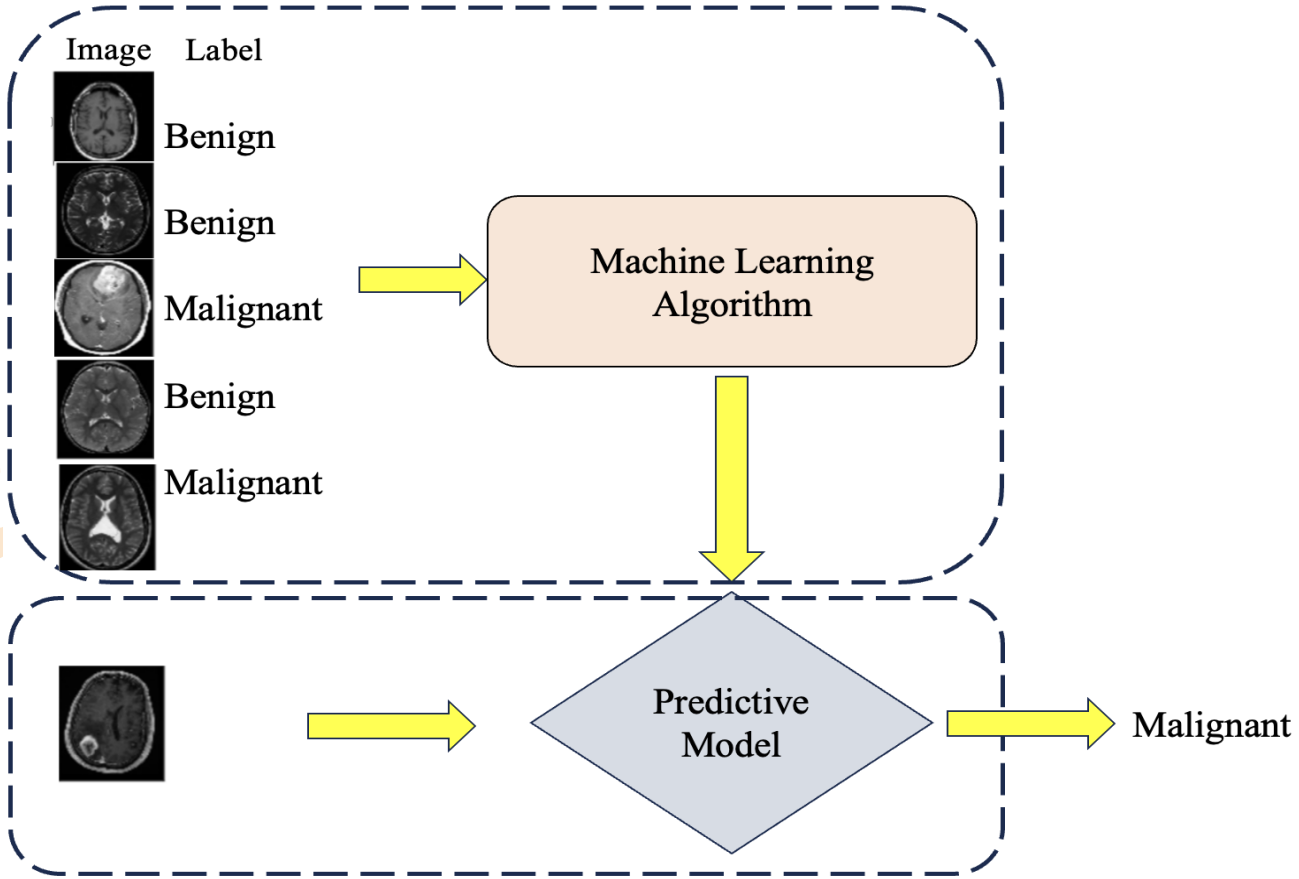
Given a labeled set of input-output pairs $\boldsymbol{D} = (\boldsymbol{x_i}, \boldsymbol{y_i})$ for i = 1 to n, the goal of Supervised learning is to learn a ***mapping function*** from $\mathbf{x_i}$ to $\mathbf{y_i}$.



Supervised Learning – Classification

Classification - Predicting the label

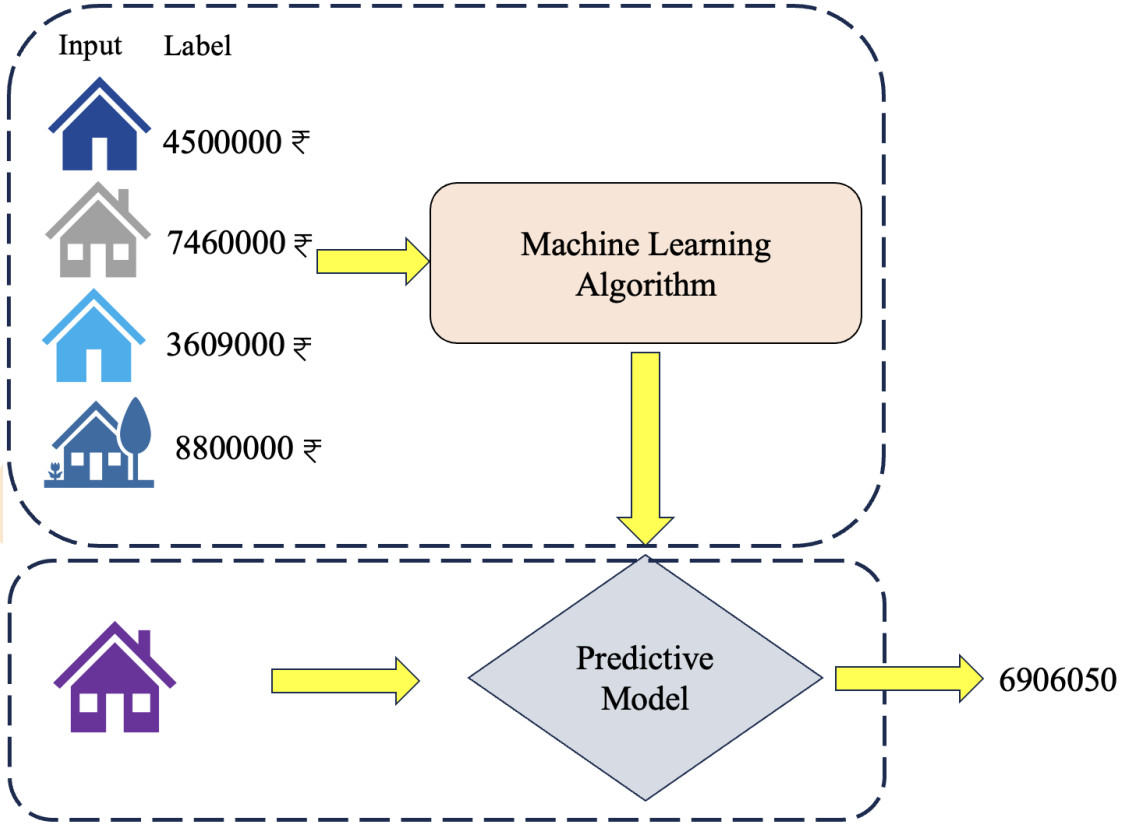
Medical diagnosis



Supervised Learning – Regression

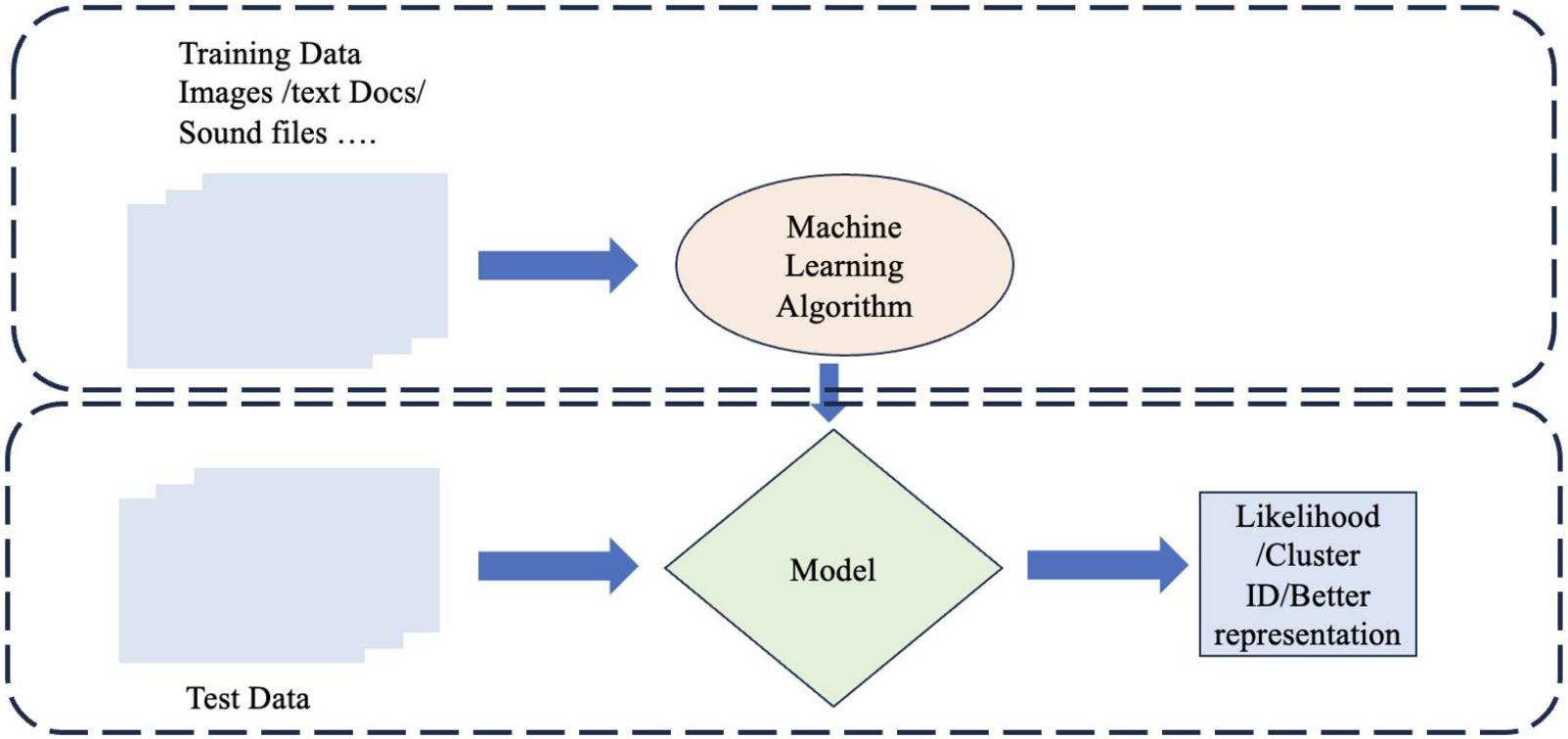
Regression – Predicting a continuous value

Selling price of house?



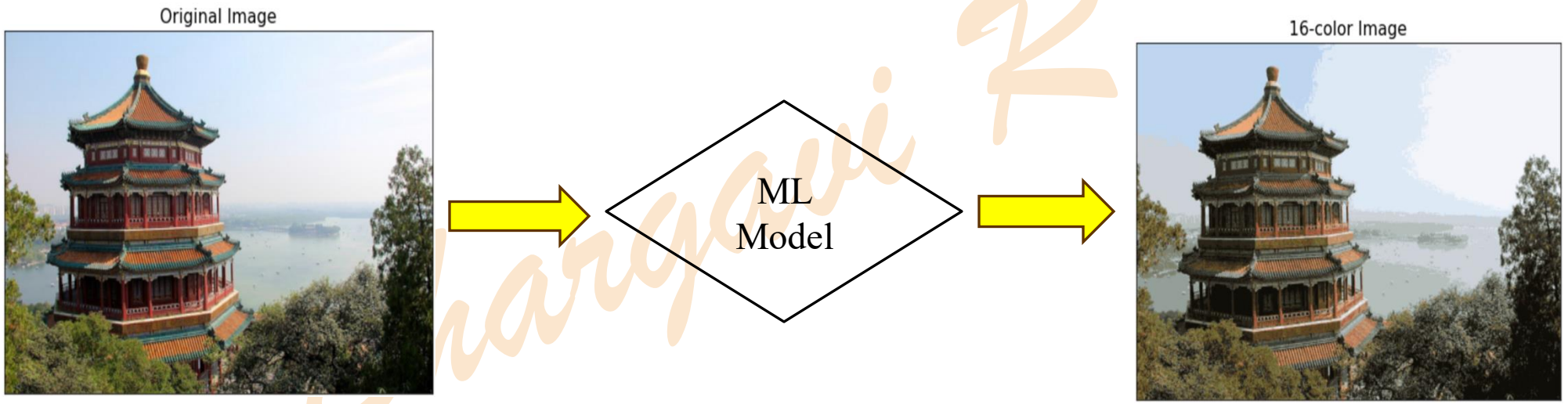
Unsupervised Learning

Given input data $\mathbf{x}_i$ for i = 1 to n, the goal of Unsupervised learning is to find "interesting patterns" from $\mathbf{x}_i$.



Unsupervised Learning - Data Compression

Color compression



Reinforcement learning

Reinforcement learning involves an agent learning to make decisions by interacting with an environment.

The agent receives feedback in the form of rewards or penalties based on its actions.

The goal is to learn a policy that maximizes the cumulative reward over time.

Parametric Models

- Assumes a functional form that summarizes the data.
- Involves learning/ estimating the parameters of the function from the training data
- **Examples** : Logistic Regression, Naïve Bayes, Perceptron etc.
- **Advantages**:
  - Simple – Easier to understand and interpret the results
  - Speed – Parametric models are fast to learn from the data
  - Do not require much data for training.
- **Disadvantages**:
  - Poor fit – Model/ Functional form assumed may not match the true unknown function
  - Constrained - By choosing a functional form these methods are highly constrained to the specified form.
  - Limited complexity – More suitable for simpler problems

Non-Parametric Models

- Do not make explicit assumptions about the functional form
- Non-Parametric models seek to best fit the training data in constructing the mapping function, whilst maintaining some ability to generalize to unseen data.
- **Examples : k-Nearest Neighbors, Decision Tree etc.**
- **Advantages:**
  - Flexible - Capable of fitting a large number of functional forms.
  - Powerful - No assumptions (or weak assumptions) about the underlying function.
  - Performance - Result in higher performance models for prediction.
- **Disadvantages**:
  - More training data required
  - Slower
  - Overfitting

Course page URL details:

https://bhargaviren.github.io/