Intro To ML & NN

21 July 2025 20:50

{Rohit Jindal} → [Sr D.S Target] [IISc Bangalore]

Civil Engineer → DTU

Mastercard $SGR

↓

ICAIF

- Logistics → [9 to 11:30 pm] → 11 pm (11:15 pm)
- Mode of Communication → (English)
- Notes and Tools
- Doubts
- Chat

→ (Dashboard) → g colab → [link] → shareable

Concept →

* [11:15 pm]

* [Intro to ML & NN] (support)

Intro to ML

Agenda

- Already Learnt
- DS ML Libraries
- Machine Learning & Classical Programming (SDE)
- Types of Problems (ML) (High Level)
  - R → Depth
  - C → Depth
  - :

(Advance) [optional]

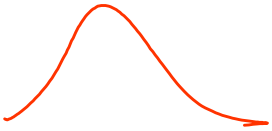
Let's Recall . . .

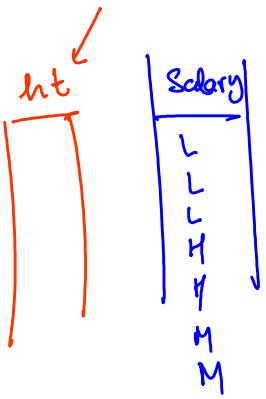
Let's briefly discuss what all we learnt in previous modules

1. **DSML Libraries** - learnt how to load and process the data
2. **Probability and Statistics** - learnt how to understand the data
3. **Coordinate Geometry & Linear Algebra** - learnt how to work with high dimensional spaces
4. **Calculus and Optimization** - learnt some optimisation techniques like GD, super useful as we learn ML

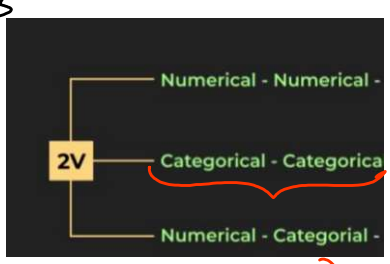
DSML libraries → Data Preprocessing + Plotting

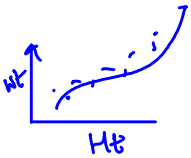
1 variable
- Categorical → Bar Plot, Count Plot
- Numerical (ht) → histogram, KDE

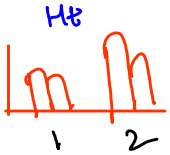




2 variables



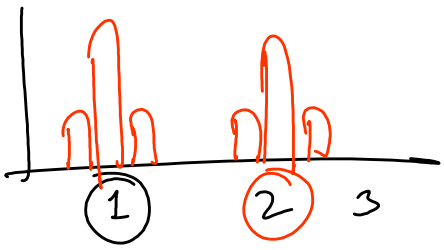




Stacked Bar, Dodged Bar

Region
1
2
3

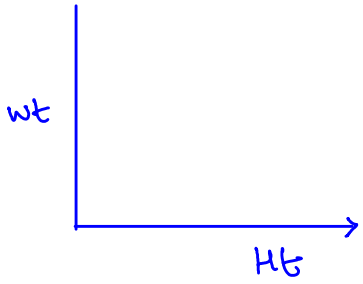
Sales → L, M, H



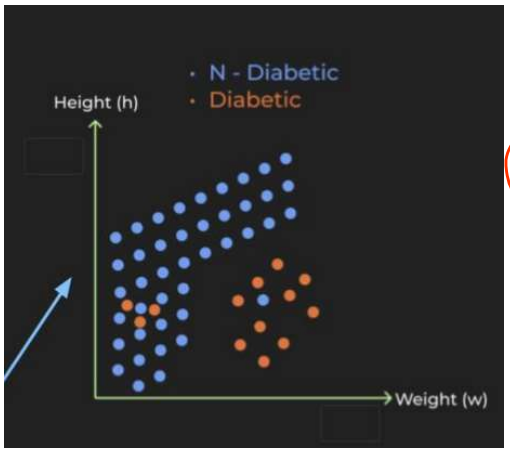
Can we Plot 3 variables ?

Can we plot 3 variable ?

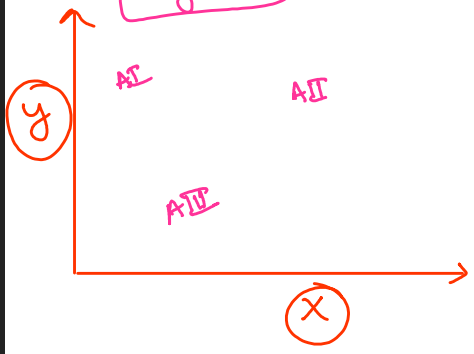
	Ht	Wt	Diabetic
①	2	3	No
②	3	5	Yes



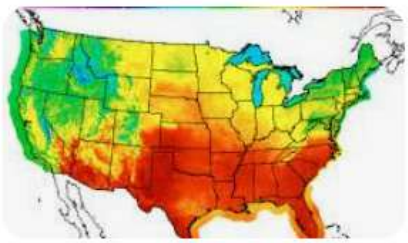
Heatmap of Correlations



Bangalore



X = Count of (Restaurant)



II Probability & Statistics

(P(diabetic))

(Conditional Probabilities)

Machine Learning is diff from Classical Programming

I ML → The field of study that makes computers capable of learning without being explicitly programmed

(I) ML → The field of study that makes computers capable of learning without being explicitly programmed

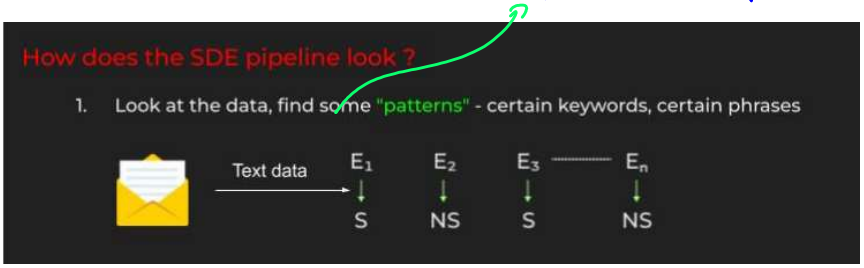
(Arthur Samuel)

Task = Human / Other way

Task → Email → spam / Not spam

Software Engineer → Powerful PC

human is explicitly finding patterns



Hardcoded rules

"Nigerian Prince" w1 "Lottery" w2 "Quick action" w3 "Credit card details" w4

[if w1 | w2 | w3 -- in email :
return spam]

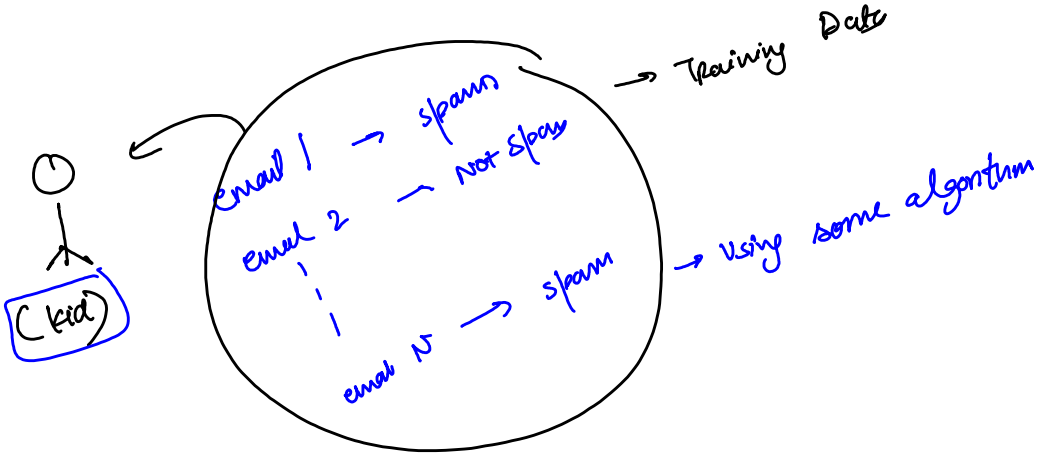
lottery → LotteRY →

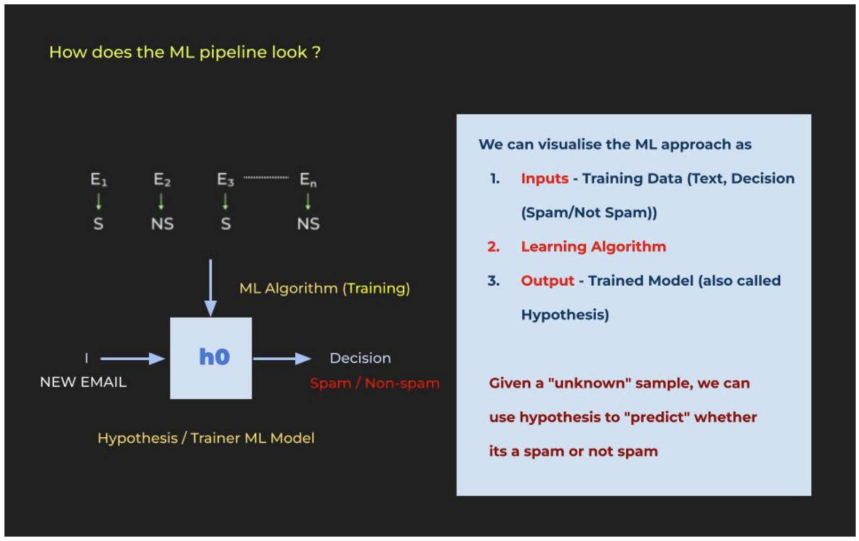
[How ML engineer will solve this problem]

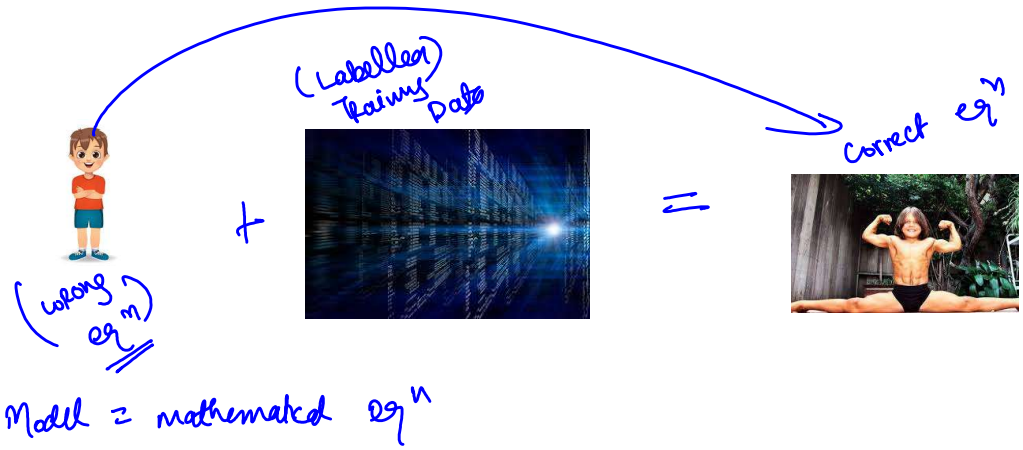
ML → I do not know rules explicitly, but I have lot of examples which are already labelled

lot of examples

→ let the machine learn general patterns







Model = mathematical eqn

(I) ML → (Tom Mitchell)

(1) ML → (Iom

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.

Algo (data) (T)

Task | Experience | Performance

(Spam / Not Spam) | [Data] | * (Any metric)

Performance (Accuracy) improves with Experience (more data)

Accuracy → 100 spam emails → Model → 80 → spam, 20 → Not spam

Accuracy = 80%

(# correct / Total) = accuracy

T	E	P

(1) [Predict Price of stock] — historical prices

(2) Categorise Image into orange, kinno, mossambi

Train using Images + Labels using a model which can "classify" an image as O, K, M	Any metric which gives a high(er) score for correct label, and low(er) score for incorrect labels

POINTS TO REMEMBER

- We use numpy and pandas for data loading, processing and manipulation
- The type of plot can be decided looking at the kind of data we have.
- In conditional probability, we check the probability of an event with respect to another event which has happened.
- If xi is present in the data, it becomes a data analysis problem, else it becomes a ML problem.

POINTS TO REMEMBER

- SDE pipeline : Provide Rules and Input, Model will give decision.
- ML Pipeline : We provide Input data and ML algorithm, algorithm will mine rules itself and decide output based on them
- Tom-Mitchell def : Study of algorithms which improves performance (P) at Task (T) with Experience (E)
- Types of task based on task - classification, regression, clustering, recommendation, forecasting

Break — 10:13 pm

Classification based on type of learning

(features) (Problem) (Label)

x_1	x_2 ----	x_d	y → category (S | M | L)

I Classification

* [Multi Class - Classification Problem]
* [Binary Class - Classification Problem]

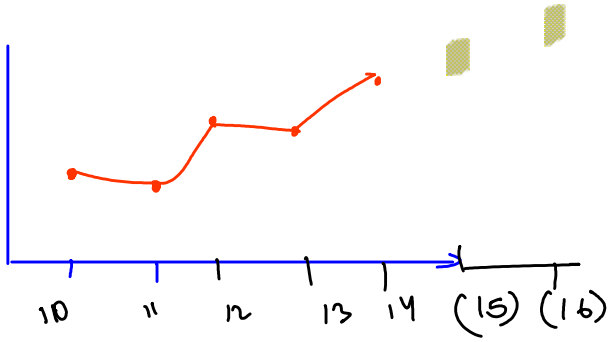
* [Binary Class – Classification Problem]
(2)
→ y is a category

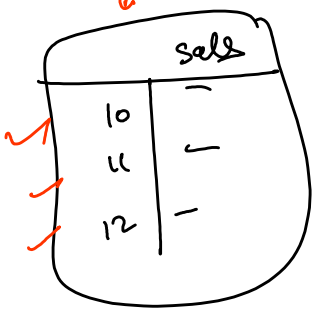
II Regression (Predict a numerical value)

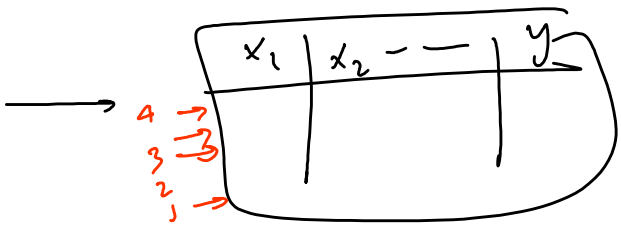
x_1	x_2	x_3	–	–	x_d	y
						300k$ 1M$

I	II
Classification: Classify into one of categories, e.g: citrus 🍊 fruits classification	**Regression :** predict a real value, e.g, price 💵prediction using stock fundamentals

III Forecasting (Time Series)

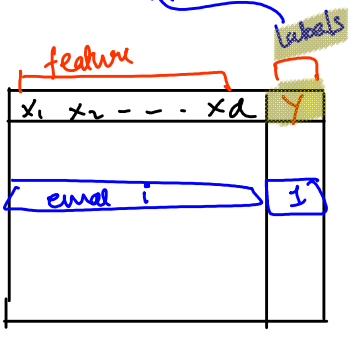




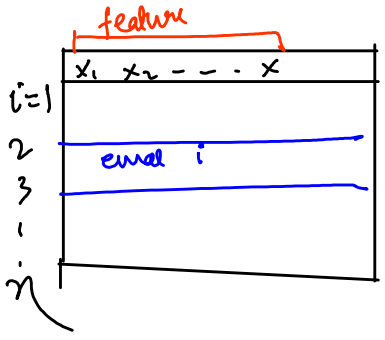


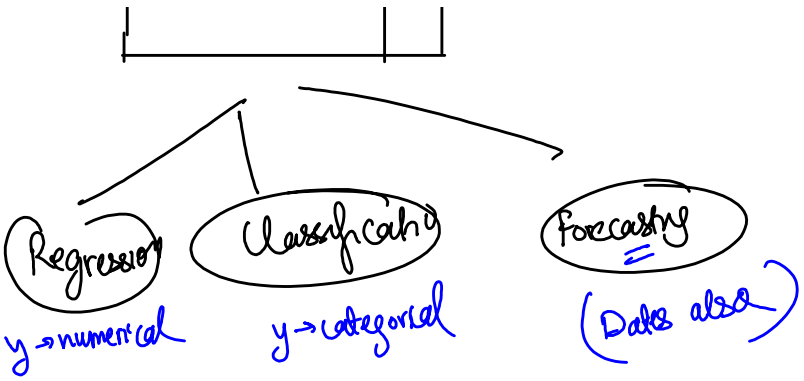
Vs

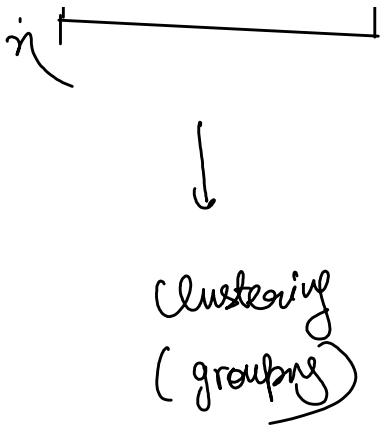
Supervised Learning



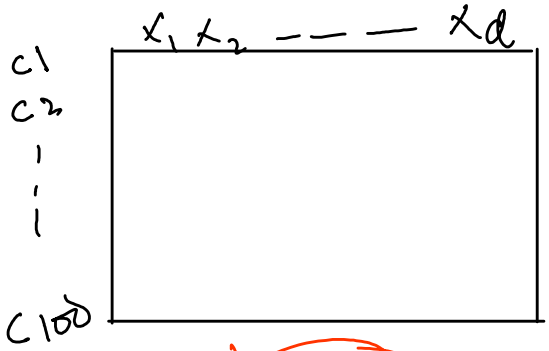
Unsupervised Learning

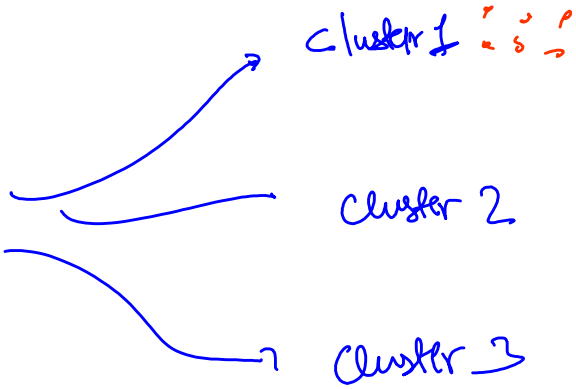


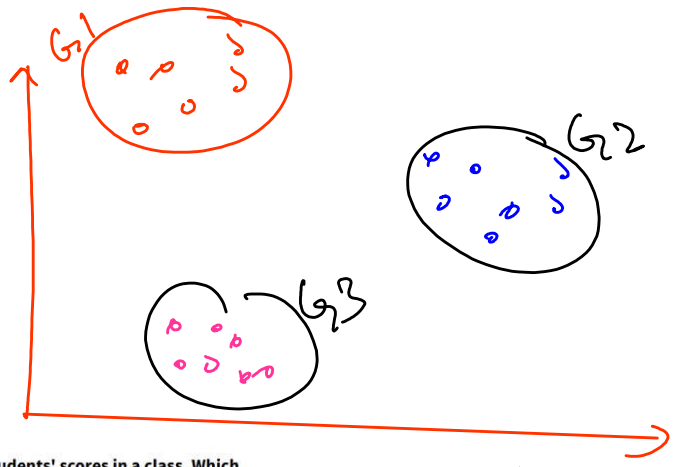




Amazon







You want to display the distribution of students' scores in a class. Which plot is best suited for this task?

4 options

Active Duration (Most preferred: 30 seconds)

Appears for: 30 Secs

- A Bar Plot
- B Histogram ✓
- C Scatter Plot ✗
- D Line Plot ✗

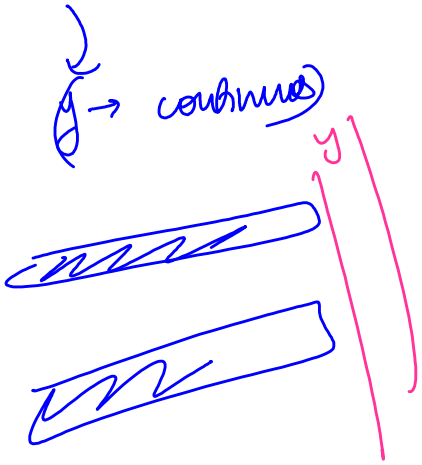
What is the main purpose of regression analysis in machine learning?

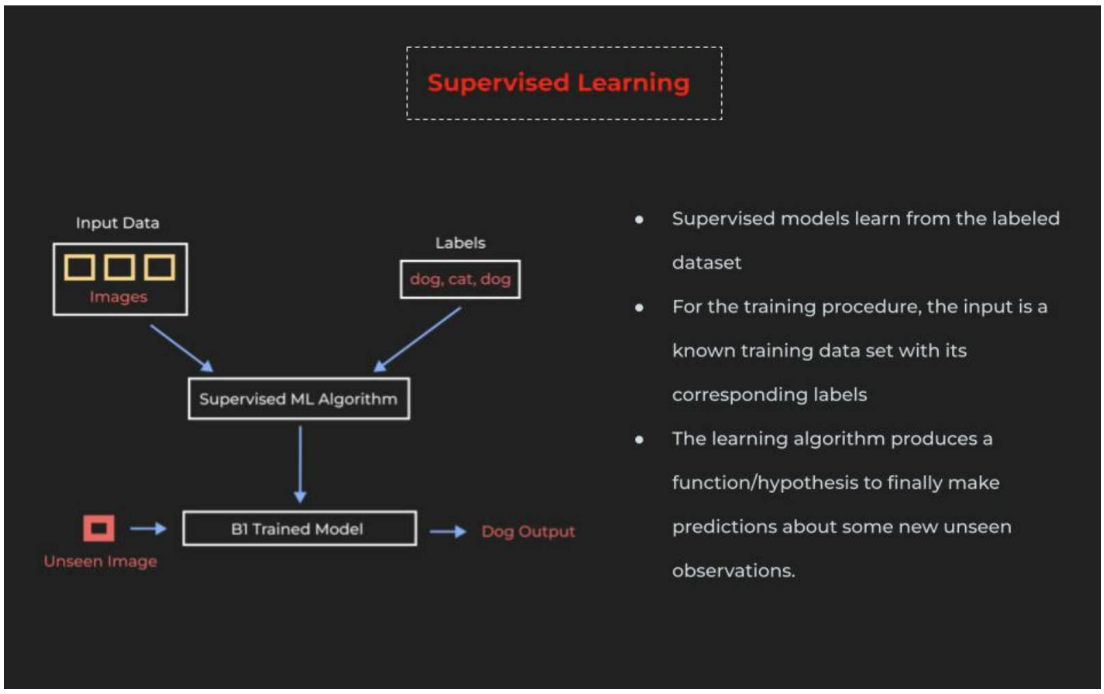
4 options

Active Duration(Most preferred: 30 seconds)

Appears for	30 Secs

A	To classify data into distinct categories.
B	To predict a continuous outcome or value.
C	To perform clustering on the dataset.
D	To visualize high-dimensional data.





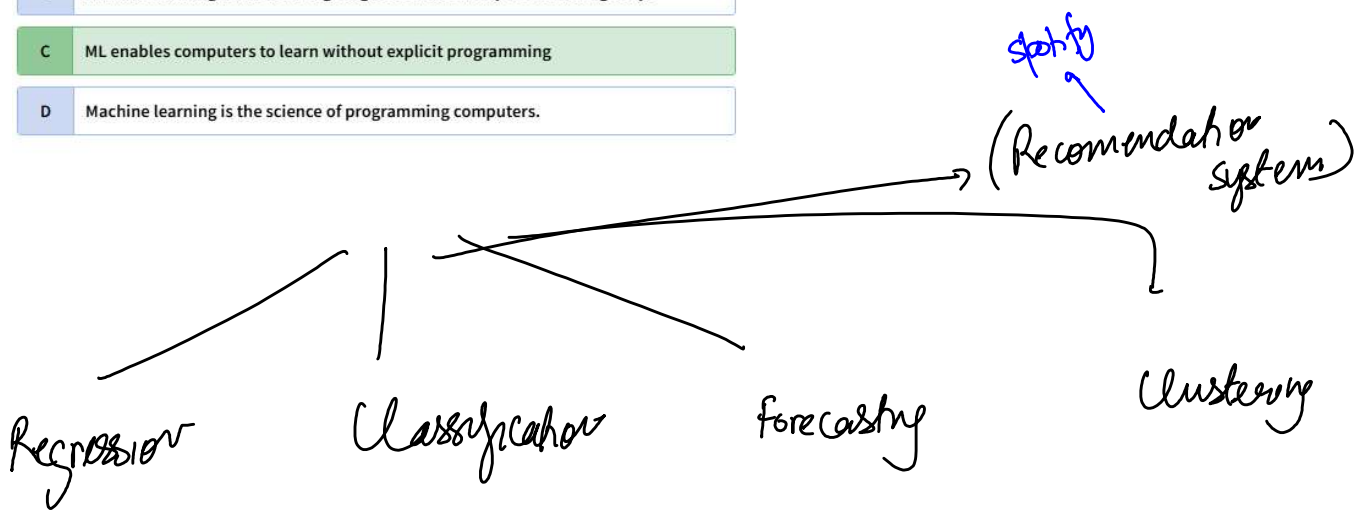
Which of the following provides the best explanation of Machine Learning?

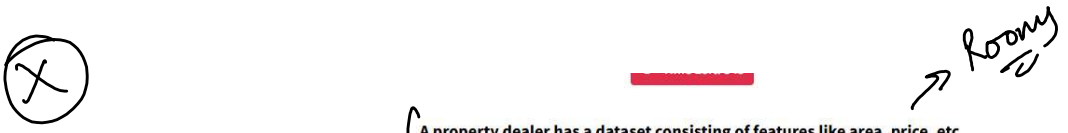
4 options

Active Duration (Most preferred: 30 seconds)

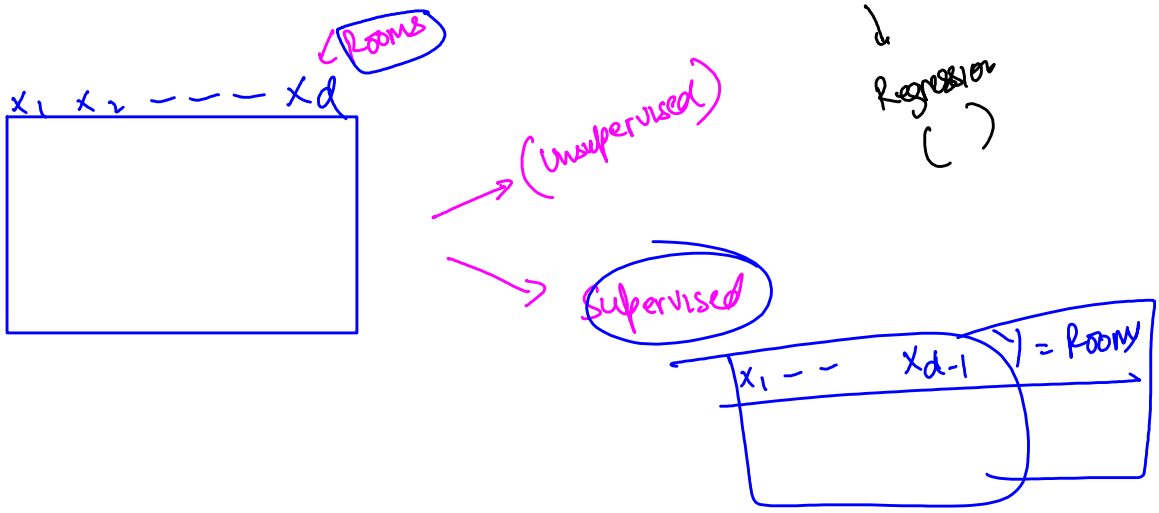
Appears for	30 Secs

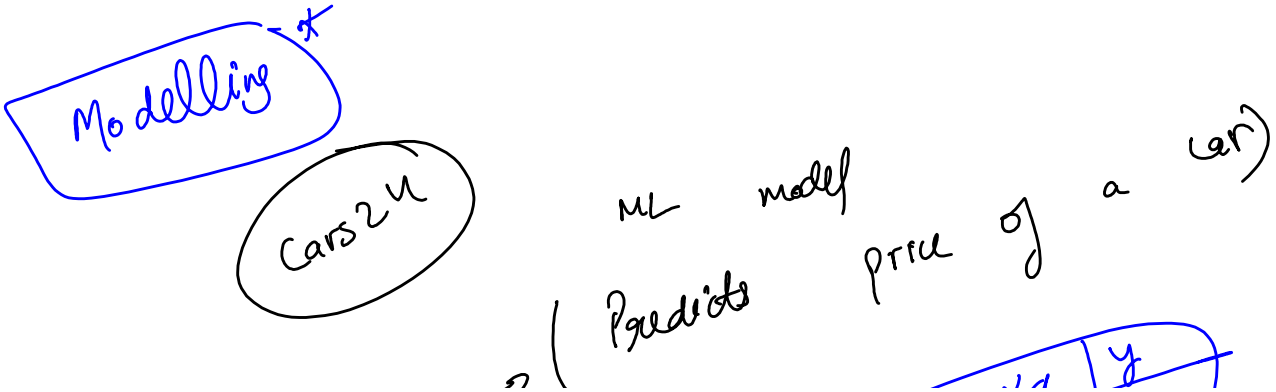
A	Machine learning learns from labelled data
B	Machine learning is the field of giving robots the ability to act intelligently.
C	ML enables computers to learn without explicit programming
D	Machine learning is the science of programming computers.

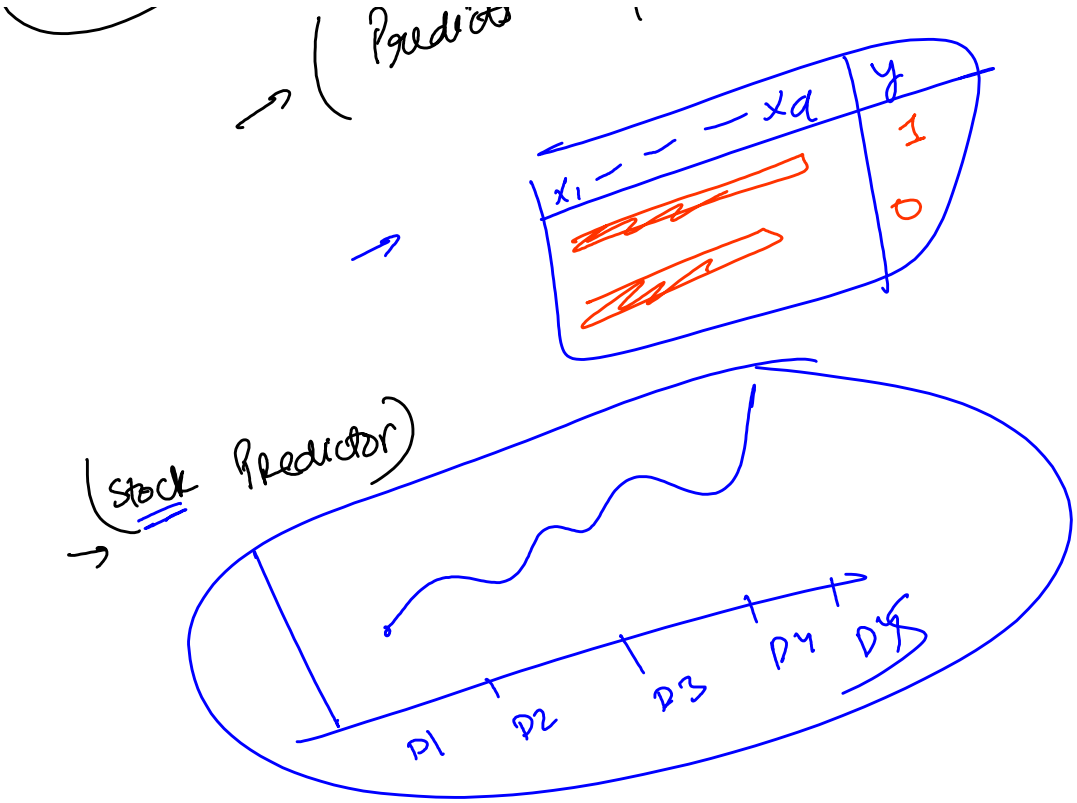




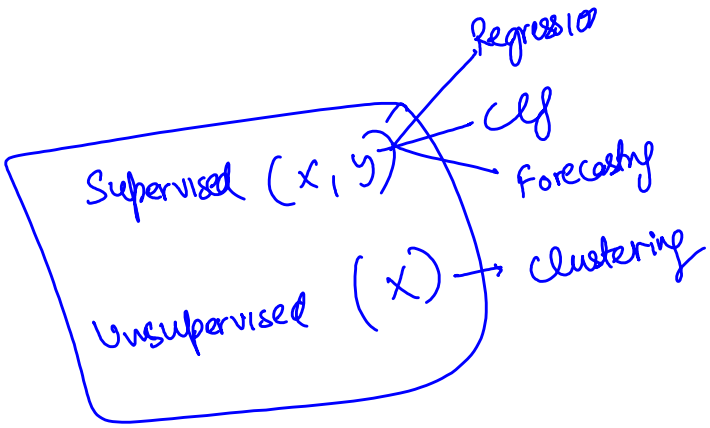
A property dealer has a dataset consisting of features like area, price, etc.
Now a customer comes to him asking for a property with a certain number of rooms. Which kind of machine learning technique should the property dealer use from the following?





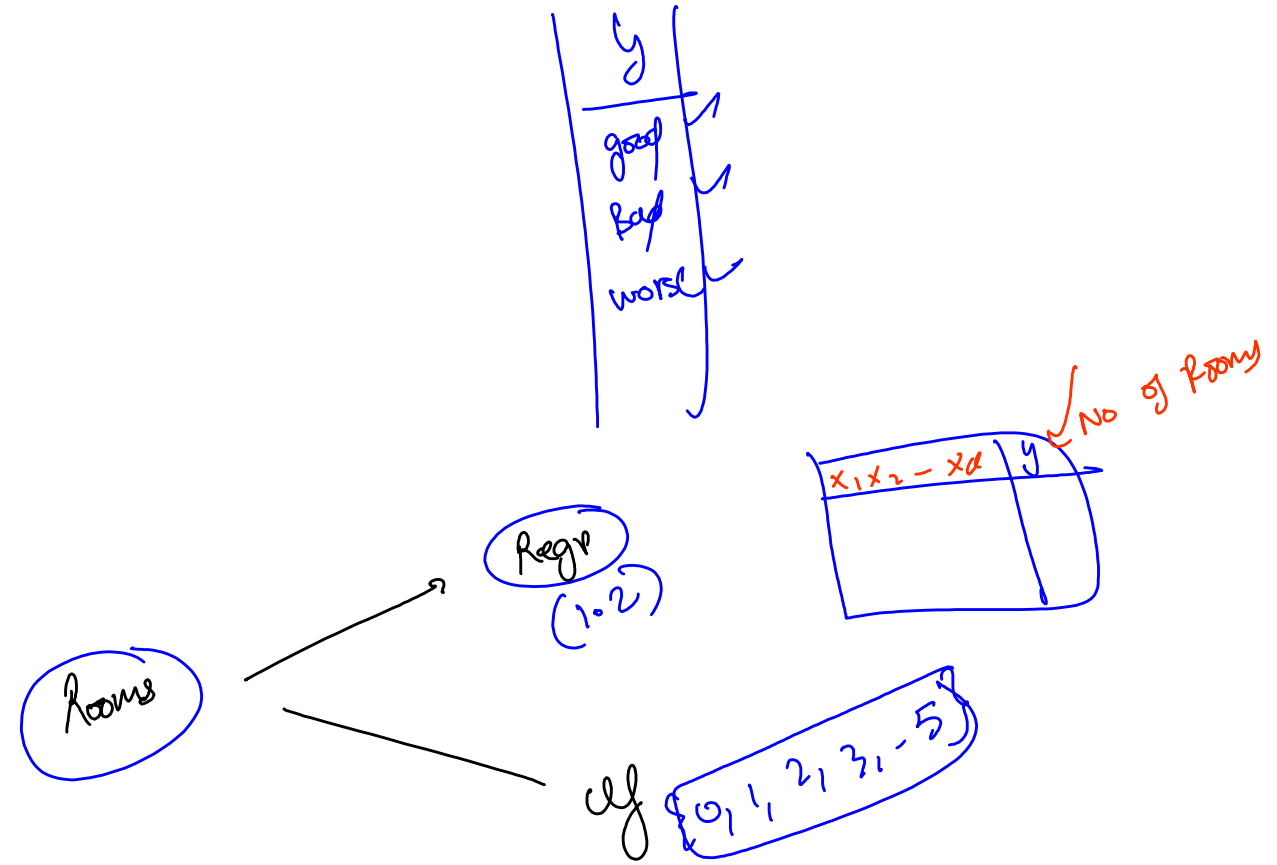


You are working on a project to predict stock prices based on historical market data, and you have a dataset with features such as past stock prices, trading volume, and other market indicators. Is this a supervised or unsupervised learning problem?

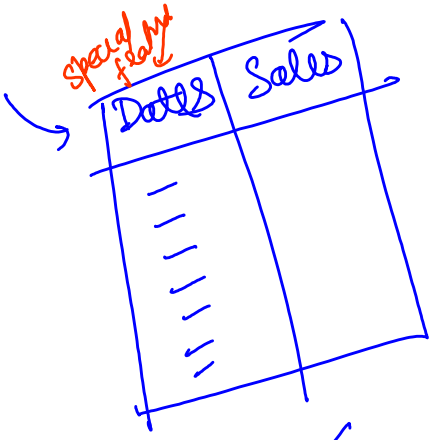


Your supervisor asks you to create a machine learning system that will help your human resources department classify jobs applicants into well-defined groups. What type of system are you more likely to recommend?

A	an unsupervised machine learning system that clusters together the best candidates.
B	you would not recommend a machine learning system for this type of project.
C	a supervised machine learning system that classifies applicants into existing groups.



Forecasting problems



Ho do you forecast unknow future factors such as Carona?