COMPUTATIONAL THINKING & AI

AI Visionaries • Innovators of India

CLASS 8

~ The Capstone Year ~

120 Pages • Master Edition • Final Book of Series

CHIRAG

Editone Kids • Editone International Pvt. Ltd.

COMPUTATIONAL THINKING & AI

AI Visionaries • Innovators of India

For Class 8 · Final Book · Series Complete!

Published by

Editone Kids · A division of

Editone International Pvt. Ltd.

Naraina Industrial Area, New Delhi

www.editone.kids

Aligned with CBSE CT & AI Curriculum 2026–27, NEP 2020, NCF–SE 2023, and DPDP Act 2023

First Edition · 2026 · Master Capstone

Welcome!

A NOTE FOR GROWN-UPS

Dear Teacher and Parent,

Class 8 is the **CAPSTONE YEAR** – the culmination of an 8-year journey. Your child moves from "specialized AI builder" to **"AI visionary & future planner"**. They master full project lifecycles, math foundations, sector-wide AI, deep ethics, and plan their career path.

What's NEW in Class 8:

- **Full 7-step AI lifecycle** mastery
- **Math behind ML** – stats, vectors, gradient descent
- **AI across ALL sectors** (14 industries!)
- **Ethics deep dive** – AGI, alignment, EU AI Act
- **Career planning** – IITs, IIITs, BITS, IISc + global
- **Competitive exams** – JEE, KVPY, IOI
- **Entrepreneurship** & startup journey
- **10-year personal plan**

The Graduation Promise:

By the end of Class 8, your child will have completed an 8-year CT/AI journey, built a portfolio, mastered math foundations, and have a clear path to college!

~110 hours | CBSE 2026–27 ready | Capstone-driven

Guide

HOW TO USE THIS BOOK

Class 8 introduces master-level sign types:

Capstone Concept

Mastery-level AI ideas.

Math Behind AI

NEW! Real math foundations!

Sector Spotlight

NEW! AI in industry!

Ethics Deep Dive

Critical moral questions!

Career Path

NEW! Plan your future!

India Page

Indian AI excellence!

8-Year Journey Complete!

This is your CAPSTONE book. Keep ALL 8 books as a treasured library — they tell your AI journey from Class 1!

Keep: Engineering Journal + Portfolio folder + 10-year plan!

Hello!

MEET THE AI VISIONARIES!

10 friends + 1 NEW = 11 AI Visionaries – capstone team!

Diya
CV Specialist!

Arjun
Architect!

Mithu
Analyst!

Deep
Multi-AI!

Tara
Project Lead!

Veer
Data Sci!

Riya
Coder!

Kabir
Detective!

Ishaan
ML Engineer!

Zara
Robotics!

NEW!

Ananya
AI Researcher – Future Prof!

Meet Ananya:

An AI researcher passionate about academics. She wants to teach at an IIT! She'll guide you through career planning & research careers!

11 AI VISIONARIES • THE GRADUATION TEAM!

WHAT'S IN THIS BOOK?

5 CHAPTERS • 120 PAGES • MASTER CAPSTONE!


CHAPTER 1
AI PROJECT MASTERY
From idea to deployment — the master's path!
PROJECT PLAN
IDEA
DESIGN
BUILD
TRAIN
DEPLOY
IMPACT

Guide

HOW TO USE THIS BOOK

Class 8 introduces master-level sign types:

Capstone Concept
Mastery-level AI ideas.

Math Behind AI
NEW! Real math foundations!

Sector Spotlight
NEW! AI in industry!

Ethics Deep Dive
Critical moral questions!

Career Path
NEW! Plan your future!

India Page
Indian AI excellence!

8-Year Journey Complete!
This is your CAPSTONE book. Keep ALL 8 books as a treasured library — they tell your AI journey from Class 1!

Keep: Engineering Journal + Portfolio folder + 10-year plan!

STEP 1
PROBLEM FRAMING

A CLEAR PROBLEM = HALF THE SOLUTION!

The CRISP framework:

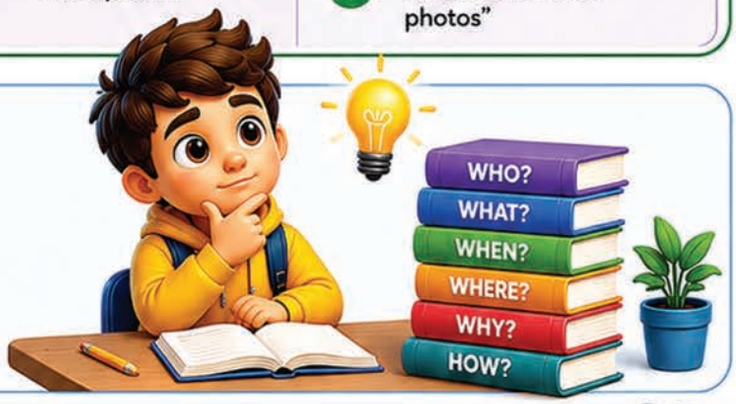
- **C** Concrete – specific, measurable
- **R** Relevant – real user need
- **I** Important – worth solving
- **S** Solvable – by AI specifically
- **P** Possible – data & resources available

Bad vs Good problem framing:

Problem	Issue	Better Version
"Make AI smarter"	Too vague	"Improve image classifier accuracy from 85% to 95%"
"Detect cancer"	Too broad	"Detect breast tumours in mammograms with 98% recall"
"Help farmers"	Not specific	"Identify 10 crop diseases from leaf photos"

Ask:

- WHO needs this?
- WHAT problem do they face?
- WHEN will they use the solution?
- WHERE will it be deployed?
- WHY does it matter?
- HOW will we measure success?



YOUR TURN: Frame a CRISP AI problem for your school!

Problem: _____

STEP 2
DATA SOURCING

DATA IS THE NEW OIL — AND THE NEW GOLD!

Where to find data:

1 Public Datasets

Kaggle, UCI, Hugging Face — millions free!

2 Indian Datasets

data.gov.in, AI4Bharat — Indian context!

3 Build your own

Camera, sensors, surveys

4 Web scraping

From public websites (ethically!)

Famous datasets:

- **ImageNet:** 14M images, 20K categories
- **MNIST:** 70K handwritten digits
- **COCO:** Common Objects in Context
- **SQuAD:** 100K Q&A pairs
- **Common Crawl:** Entire web text!
- **AI4Bharat IndicCorp:** 600B+ Indian language tokens!

DATA ETHICS RULES:

- Get permission if personal data
- Follow DPDP Act 2023 (India)
- No private images without consent
- No copyright-protected content
- Diverse, representative data

STEP 2B
DATA CLEANING

REAL DATA IS MESSY — CLEAN IT FIRST!

Common data problems:

Problem	Example	Fix
Missing values	Age: NULL	Fill with mean or drop
Duplicates	Same row 3 times	Remove duplicates
Wrong types	"25" instead of 25	Convert types "25" → 25
Outliers	Age: 200 yrs!	Investigate & fix
Inconsistent	"Mumbai", "MUMBAI", "mumbai"	Standardize A → A
Typos	"Banglore" (Bengaluru)	Correct or remove Banglore → Bengaluru

Python data cleaning

```python
# Python data cleaning
import pandas as pd
df = pd.read_csv("students.csv")  # 1. Remove duplicates
df = df.drop_duplicates()
# 2. Fill missing ages with mean
df["age"] = df["age"].mean(), inplace=True
# 3. Standardize city names
df["city"] = df["city"].str.title()
# 4. Remove age outliers
df = df[(df["age"] > 0) & (df["age"] < 120)]
```

Engineers spend 60–80% of time on data cleaning!

STEP 3

FEATURE ENGINEERING

PICK & CREATE THE RIGHT INPUTS FOR AI!

Big Word: FEATURE ENGINEERING

Transforming raw data into **MEANINGFUL** inputs for AI!
This is where **ART** meets science!

Examples:

Raw Data	Engineered Feature	Why Useful
Birthdate "2010-03-15"	Age = 15	Easier to use!
Address "Mumbai, MH"	City="Mumbai", State="MH"	Separate features
Date "Sat 17/Jan/2026"	Day_of_week=Sat, Month=Jan	Cycles matter!
"Loved this book!"	word_count=4, has_positive=1	Numeric for AI
Price = 999	Price_tier = "Mid"	Categories!

Feature engineering techniques:

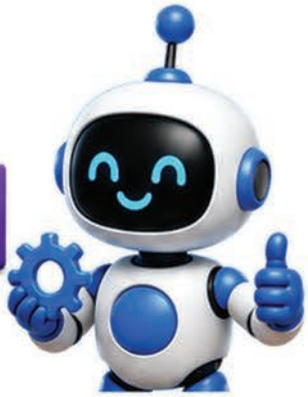
- **Encoding:** Convert categories to numbers
- **Scaling:** Same range (0–1)
- **Binning:** Group ages into youth/adult/senior
- **Polynomial:** Create X^2 for non-linear patterns
- **Datetime:** Extract year/month/day/hour

PRO TIP

Good features > fancy model!
Andrew Ng (Stanford AI): "Applied ML is **feature engineering**"!

STEP 4

MODEL SELECTION



RIGHT MODEL FOR THE RIGHT PROBLEM!

Common ML models & when to use:

Model	Best For	Difficulty
Linear Regression	Predict numbers (price, temp)	★
Logistic Regression	Yes/No classification	★★
Decision Tree	Interpretable rules	★★
Random Forest	Tabular data, robust	★★★
K-Means	Clustering (no labels!)	★★
Neural Network	Complex patterns	★★★★
CNN	Images (Computer Vision)	★★★★
Transformer	Language (LLMs!)	★★★★★

The model selection method:

1. Start SIMPLE (Linear Regression!)
2. Establish baseline accuracy
3. Try more complex if simple isn't enough
4. Compare 3–4 models on same data
5. Pick best balance of accuracy + speed

Sometimes Linear Regression beats Neural Network —
try simple first!

STEP 5
EVALUATION METRICS

"MY MODEL IS 90% ACCURATE!" – IS THAT GOOD?

The accuracy trap:

If 99% of emails are NOT spam, a model that says "not spam" always = 99% accurate! **Useless!**

Better metrics:

ACCURACY = (Correct) / (Total)

PRECISION = (True Positives) / (True+False Positives)

"Of all things I predicted YES, how many were really YES?"

RECALL = (True Positives) / (True Positives + False Negatives)

"Of all real YES, how many did I catch?"

F1 SCORE = 2 × (Precision × Recall) / (Precision + Recall)

When to prioritize each:

- **Precision:** Spam filter (don't mark good email as spam!)
- **Recall:** Cancer detection (don't MISS any case!)
- **F1** Balanced – most general use

MEDICAL AI

Cancer screening prioritizes RECALL (97%+) over precision.
Better to falsely flag a healthy person than miss a real cancer!

INDIA PAGE

ISRO's AI PROJECTS!

HOW ISRO USES THE 7-STEP LIFECYCLE!

1 Bhuvan AI

Satellite image analysis – Forest cover, fires, floods, crop monitoring!

2 Chandrayaan-3

Autonomous landing AI! Pragyan rover's SLAM navigation!

3 Aditya-L1

India's sun mission – AI detects solar flares!

4 SCATSAT-1

Ocean & weather AI predictions!

5 NaviC AI

India's own GPS – ML for accuracy!

6 Mission Mausam

2026: AI weather forecasting for farmers!

ISRO + AI

ISRO planning AGNIBAAN reusable rocket with AI autopilot!
Indian Space Station with AI-controlled subsystems by 2035!

Career path: ISRO hires 500+ engineers/year! Need: AI, robotics, aerospace skills!

My ISRO dream project: _____

STEP 6 DEPLOYMENT

FROM YOUR LAPTOP → MILLIONS OF USERS!

Deployment options:

Type	Best For	Tools
Mobile app	Consumer apps	TensorFlow Lite, CoreML
Web app	Wide reach	Flask, FastAPI, Streamlit
Cloud API	Scale to millions	AWS SageMaker, GCP AI
Edge device	IoT, embedded	Raspberry Pi, Jetson
Browser	Demos, education	TensorFlow.js

Deployment challenges:

- **Latency:** Response time < 1 second
- **Scale:** Handle 1M+ users
- **Security:** No data leaks!
- **Size:** Compress 1GB model to 50MB
- **Offline:** Work without internet

AAROGYA SETU

India's COVID app: 200M+ users! Deployed across iOS, Android, KaiOS for feature phones! Privacy-preserving AI!

"ML Engineer" jobs focus heavily on deployment – **high-paying career!**

STEP 7 MONITORING

MODELS DEGRADE OVER TIME — MONITOR & FIX!

BIG WORD: MODEL DRIFT

When model accuracy **DROPS** over time because real-world data changes!

EXAMPLE: SPAM FILTER DRIFT

YEAR	ACCURACY	REASON
2024 (launch)	97%	Trained on 2024 spam
2025	92%	New scam types
2026	85%	AI-generated spam
2026 (retrained!)	96%	Updated with new data

WHAT TO MONITOR:

- Accuracy/ Precision/Recall trends
- Response times
- Error rates
- User feedback signals
- Data distribution shifts

THE MLOPS CYCLE:

Deploy → Monitor → Detect drift → Collect new data → Retrain → Deploy ...

MLOPS ENGINEER

Top-paying job (₹25–60 LPA)! Combines ML + DevOps. Big demand in India.

CASE STUDY

AAROGYA SETU

INDIA'S MOST-USED AI APP — ANALYZED!

Aarogya Setu (2020–2024):

- **200 MILLION** users in 1 month!
- Available in **12 Indian languages**
- **Bluetooth**-based proximity detection
- **COVID** risk assessment via AI

THE 7 STEPS IN AAROGYA SETU:

STEP	WHAT THEY DID	CHALLENGE
1 Problem	Track COVID exposure	Privacy concerns
2 Data	Bluetooth + GPS	Battery drain
3 Features	Distance, duration	Indoor accuracy
4 Model	Risk scoring AI	False positives
5 Eval	Test in cities	Edge cases
6 Deploy	Android + iOS + KaiOS	Scale to 200M!
7 Monitor	Track exposures	Updates daily

LESSONS LEARNED

1. Build for **ALL** phones (feature phones too!)
2. **Privacy** by design!
3. **Multilingual** from day 1!
4. Iterate fast based on **user feedback**!

CASE STUDY

AAROGYA SETU

INDIA'S MOST-USED AI APP – ANALYZED!

Aarogya Setu (2020–2024):

- **200 MILLION** users in 1 month!
- Available in **12 Indian languages**
- **Bluetooth**-based proximity detection
- **COVID** risk assessment via AI

THE 7 STEPS IN AAROGYA SETU:

STEP	WHAT THEY DID	CHALLENGE
1 Problem	Track COVID exposure	Privacy concerns
2 Data	Bluetooth + GPS	Battery drain
3 Features	Distance, duration	Indoor accuracy
4 Model	Risk scoring AI	False positives
5 Eval	Test in cities	Edge cases
6 Deploy	Android + iOS + KaiOS	Scale to 200M!
7 Monitor	Track exposures	Updates daily

★ LESSONS LEARNED

1. Build for **ALL** phones (feature phones too!)
2. **Privacy** by design!
3. **Multilingual** from day 1!
4. Iterate fast based on **user feedback**!

HYPERPARAMETER TUNING

MODEL HAS "KNOBS" TO TUNE — FIND BEST SETTINGS!

BIG WORD: HYPERPARAMETER

Settings YOU choose BEFORE training (not learned by AI).
Like recipe quantities!

COMMON HYPERPARAMETERS

HYPERPARAMETER	EXAMPLE	EFFECT
Learning rate	0.001, 0.01, 0.1	Speed of learning
Batch size	32, 64, 128	Samples per step
Epochs	10, 50, 100	Training iterations
Hidden layers	2, 5, 10	Model depth
K (KNN)	3, 5, 11	Neighbour count

TUNING STRATEGIES

Grid Search:
Try ALL combinations
(slow but thorough!)

Random Search:
Random combinations
(faster!)

Bayesian Opt:
Smart search
(modern!)

Automated:
AutoML tools
(Google AutoML!)

TIME COST

Hyperparameter tuning can take **HOURS** or **DAYS!**
Real engineers often pay **$1000s** on cloud!

COMMON PITFALLS!

LEARN FROM MISTAKES — DON'T REPEAT THEM!

TOP 10 AI PROJECT MISTAKES:

1. **Bad problem framing** – vague goals
2. **Insufficient data** – <1000 samples
3. **Biased data** – only one demographic
4. **Train/test leakage** – same data both!
5. **Wrong metric** – accuracy when need recall
6. **Overfitting** – model memorized data
7. **Underfitting** – model too simple
8. **Ignoring deployment** – model never used
9. **No monitoring** – silent failures
10. **No ethics review** – harmful outcomes

REAL DISASTER: APPLE CARD BIAS (2019)

Apple Card gave MEN **10x** higher credit limits than women with same income!

AI learned old bias from data.
Major scandal!

LESSON

Audit your data for **bias**
BEFORE deploying!
Diverse team = **better AI!**

CAPSTONE PROJECT FULL LIFECYCLE!

⭐ CAPSTONE MINI-PROJECT ⭐

Mission: Plan an AI project using all 7 steps – for a real problem!

Your Capstone:

1. **PROBLEM (CRISP):** _____
2. **DATA source:** _____ Samples needed: _____
3. **FEATURES (3):** _____ , _____ , _____
4. **MODEL chosen:** _____ Why? _____
5. **METRICS:** ☐ Accuracy ☐ Precision ☐ Recall ☐ F1
6. **DEPLOYMENT:** ☐ App ☐ Web ☐ API ☐ Edge
7. **MONITORING:** Track _____ over time

Ethics check:

Privacy: ☐ DPDP compliant ☐ Anonymized data

Bias check: ☐ Diverse data ☐ Test groups

Timeline:

Total weeks: _____ Team size: _____

This is your **CAPSTONE** – biggest project of 8 years!

BRAIN TEASERS LIFECYCLE!

LIFECYCLE BRAIN BUSTER!

1

Puzzle 1: How many steps in the AI lifecycle?

Answer: _____

2

Puzzle 2 • Metric math:

A model predicts 100 cancers. 90 real, 10 false. Real cancers were 95. Calculate:

Precision = TP/(TP+FP) = _____/(_____+_____) = _____

Recall = TP/(TP+FN) = _____/(_____+_____) = _____

3

Puzzle 3 • 99% accuracy on imbalanced data:

Spam: 1%, Normal: 99%. Model always says "Normal". Accuracy?

Answer: _____% Useful? ☐ Yes ☐ No

4

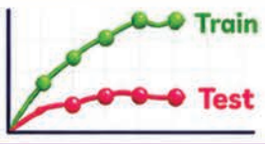
Puzzle 4 • Drift: Model launched 2024 had 95% acc. 2026 = 82%. Should we:

☐ Delete model ☐ Retrain on 2026 data ☐ Ignore

5

Puzzle 5 • Overfit: Train acc 99%, Test acc 60%. Diagnosis?

Answer: _____



6

Puzzle 6 • Indian app: Aarogya Setu reached how many users?

Answer: _____ million!

BEST PRACTICES SUMMARY

THE AI ENGINEER'S GOLDEN RULES!

10 COMMANDMENTS:

1. Frame the problem **CRISP**-ly first!
2. Spend **80%** time on **data**, 20% on model
3. Always have **train/val/test** splits
4. Use the right **metric** for the problem
5. Test **bias** across demographics
6. Start simple – **Linear** before Neural
7. **Document** everything (data, code, model)
8. **Monitor** in production!
9. Get **ethics** review **BEFORE** deployment
10. **Iterate** – first version won't be perfect!

THE GROWTH MINDSET:

Every project teaches you.
Failures are stepping stones.
Even Sundar Pichai's first project failed!

SCHOOL VS REAL

School: 1 perfect attempt.
Real engineering: 10 iterations, each improving!
Embrace iteration!

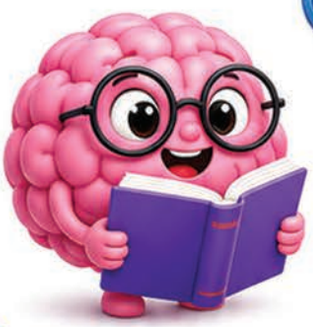
Write these 10 rules on your wall!
Read daily during your AI journey!

TRY IT YOURSELF!

LIFECYCLE MASTERY!

1. **Capstone Plan:** Complete the project plan on page 22 in detail!
2. **Dataset Hunt:** Browse Kaggle, find a dataset that interests you!
3. **Aarogya Analysis:** Pick another Indian AI app – apply 7-step analysis!
4. **Metric Practice:** Calculate precision/recall for 5 different scenarios!
5. **Bias Audit:** Pick a public AI tool. Test it on diverse inputs!
6. **Deployment Demo:** Try Google Colab to deploy a simple model!

IN CHAPTER 1 YOU MASTERED:



- The **7-step** AI lifecycle
- **CRISP** problem framing
- **Data** sourcing & cleaning
- **Feature** engineering
- **Model** selection (8+ types!)
- **Metrics:** accuracy, precision, recall, F1
- **Validation** & cross-validation
- **Hyperparameter** tuning
- **Deployment** & monitoring
- **Real case:** Aarogya Setu!

AI Project Master!

NEXT: MATH BEHIND ML!

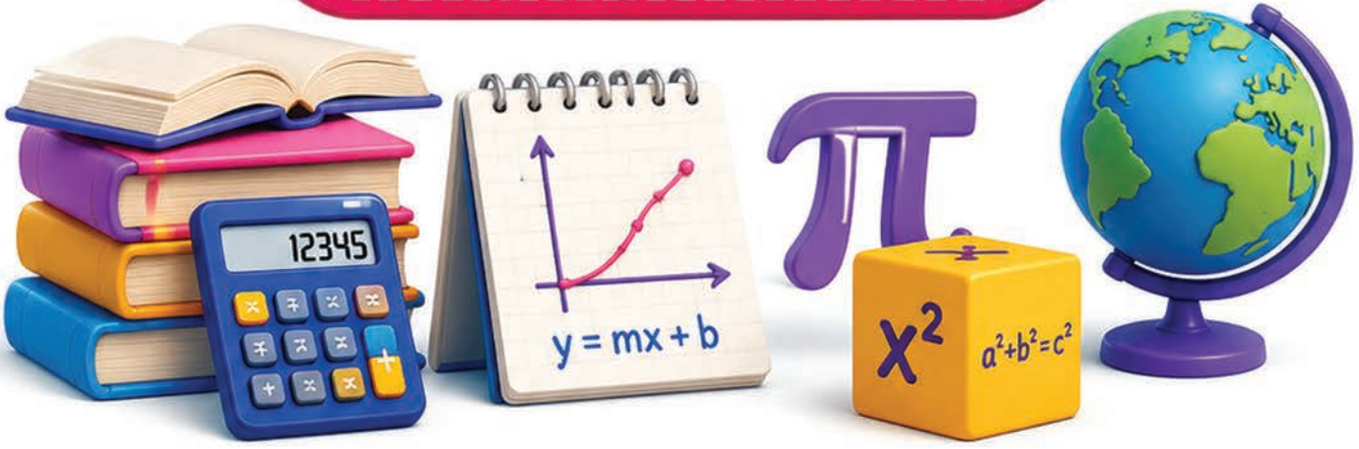
CHAPTER 2

— THE —

FOUNDATIONS

MATH BEHIND ML

From Aryabhata to AI —
the math that powers it!



WHY MATH MATTERS IN AI

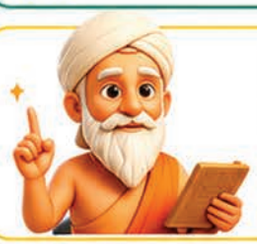
MATH IS THE LANGUAGE AI SPEAKS!

MATH YOU NEED FOR AI:

- **Statistics** – describe data
- **Probability** – handle uncertainty
- **Linear Algebra** – vectors & matrices
- **Calculus** – rates of change
- **Optimization** – find best answer

Don't worry — you'll learn it gradually!

Math Topic	Used For	When Learned
Statistics	Data analysis	Class 8–10
Probability	Prediction	Class 9–11
Algebra	Linear regression	Class 8–9
Vectors	Embeddings	Class 11–12
Matrices	Neural networks	Class 11–12
Calculus	Gradient descent	Class 11–college



ARYABHATA

India's ancient mathematician (5th century AD)!
Invented zero, place value, π approximation.
Foundation for modern AI math!

Strong math = strong AI engineer!
Don't skip!

STATISTICS RECAP

3 ESSENTIAL STATISTICS IN AI!

MEAN (average) = Sum of values / Number of values

VARIANCE = Average of (each value - mean)²

STANDARD DEVIATION = √Variance

Example: 5 students' marks (out of 10):

Step	Calculation	Result
Marks	7, 8, 6, 9, 5	–
Sum	7+8+6+9+5	35
Mean	35/5	7
Deviations	0, 1, -1, 2, -2	–
Squared deviations	0, 1, 1, 4, 4	–
Variance	(0+1+1+4+4)/5	2.0
Std Deviation	√2.0	~1.41

WHAT THEY TELL US:

- **Mean:** typical value
- **Std Dev:** spread of data
- **Small SD** → values close to mean
- **Large SD** → values spread out

Class average could be 50% with SD=5 (**consistent**)
or SD=25 (some failed, **some excelled!**)

PROBABILITY INTRO

AI PREDICTS WITH PROBABILITIES!

PROBABILITY = (Favorable outcomes) / (Total outcomes)

Examples:

Event	Calculation	Probability
Coin → Heads	1/2	0.5 (50%)
Dice → 6	1/6	0.167 (16.7%)
Card → Ace	4/52	0.077 (7.7%)
Birthday in Jan	31/365	0.085 (8.5%)

PROBABILITY IN AI:

- "Dog: 0.92" → AI is 92% sure it's a dog
- Self-driving: "Pedestrian: 0.05" → low chance of pedestrian
- Spam filter: "Spam: 0.85" → likely spam
- ChatGPT: predicts NEXT word probability

PROBABILITY RANGE:

ALWAYS between 0 (impossible) and 1 (certain).
All probabilities of all outcomes ADD to 1!

SOFTMAX

When AI outputs probabilities for multiple classes (cat 60%, dog 30%, fox 10%), it uses "softmax" to ensure they sum to 100%!

CONDITIONAL PROBABILITY & BAYES

"GIVEN SOMETHING HAPPENED, WHAT'S NEXT?"

Big Word: CONDITIONAL PROBABILITY

P(A | B) = "Probability of A given B happened"

Example:

P(rain TODAY | dark clouds) = 0.85

P(rain TODAY | clear sky) = 0.05

BAYES' THEOREM

The most famous probability formula in AI!

$$P(A \mid B) = P(B \mid A) \times P(A) / P(B)$$

Real example: Medical test

Disease affects 1% of people. Test is 99% accurate.
If you test POSITIVE, what's your chance of disease?

Math	Value	Meaning
P(disease)	0.01	1% population
P(+test | disease)	0.99	99% accurate
P(+test | no disease)	0.01	1% false positive

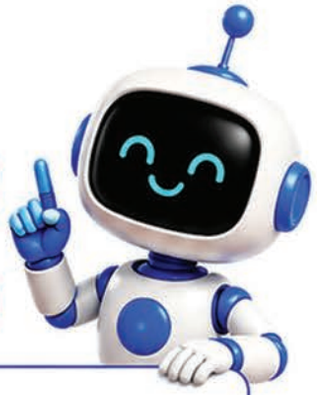
Surprise result:

P(disease | +test) = only ~50%!

BAYES IS EVERYWHERE

Spam filters, medical AI, autocorrect, weather forecast all use **Bayes**!
Even **Naive Bayes** ML model!

VECTORS
DATA AS ARROWS!



A VECTOR IS JUST A LIST OF NUMBERS!

Big Word: VECTOR

Ordered list of numbers representing a point in space!

Examples:

```
# 2D vectors (points in plane)
diya_position = [3, 5]        # x=3, y=5
arjun_position = [7, 2]       # x=7, y=2
# 3D vectors (points in 3D space)
moon_position = [384400, 0, 0] # km from Earth
# Word embedding (300D vector)
king = [0.8, 0.2, 0.9, 0.1, ...]
```

Vector operations AI uses:

- **Addition:** [1,2] + [3,4] = [4,6]
- **Multiplication by number:** 2×[1,2] = [2,4]
- **Dot product:** [1,2]·[3,4] = 1×3 + 2×4 = 11
- **Length (magnitude):** [3,4] = √(3² + 4²) = 5

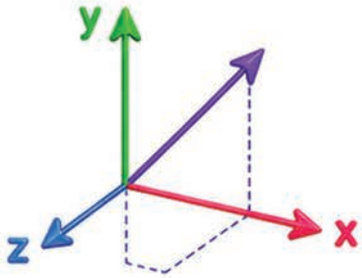
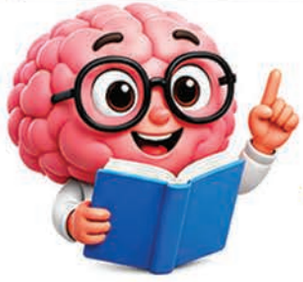


IMAGE AS VECTOR

A 28×28 grayscale image = **784-dimensional vector**!
Each pixel is one dimension!

→ [0.1 | 0.5 | 0.3 | ... | 0.9 | 0.2]

784 numbers in a list!



In AI, vectors help computers understand, compare and find patterns in data!

MATRICES

TABLES OF NUMBERS!

A MATRIX = ROWS OF VECTORS!

Example matrix:

```
# 3x3 matrix M = [[1, 2, 3], [4, 5, 6], [7, 8, 9]]
```

1	2	3
4	5	6
7	8	9

Images ARE matrices!

Grayscale 4×4 image:

```
[[0, 100, 200, 255], [50, 100, 180, 255],
[80, 120, 160, 200], [255, 255, 255, 255]]
```

0 = black, **255** = white, in between = grays!

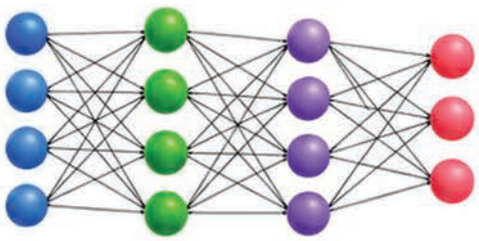
0	100	200	255
50	100	180	255
80	120	160	200
255	255	255	255

Colour images = 3 matrices stacked (R, G, B)!

HD photo = 1920 × 1080 × 3 = 6,220,800 numbers!

Matrix in AI:

- Neural network weights = matrices
- Multiply matrices = forward pass
- Image processing = matrix ops
- Linear regression = matrix math



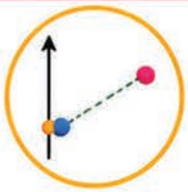
np

NumPy library makes matrix math **FAST** in **Python!**

DISTANCE METRICS

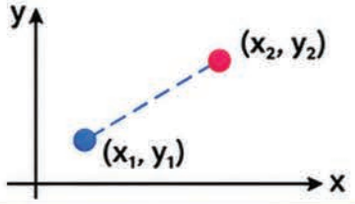
HOW "SIMILAR" ARE TWO DATA POINTS? MEASURE DISTANCE!

3 main distance metrics:



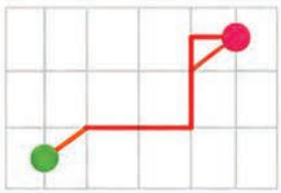
1 EUCLIDEAN distance (straight-line):

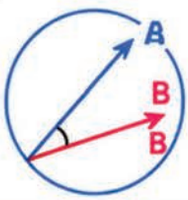
$$d = \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]}$$



2 MANHATTAN distance (grid path):

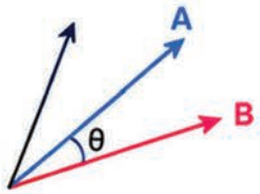
$$d = |x_2 - x_1| + |y_2 - y_1|$$





3 COSINE similarity (angle):

$$sim = \frac{(A \cdot B)}{|A| \times |B|}$$



Example: Diya at (3,5), Arjun at (7,2)

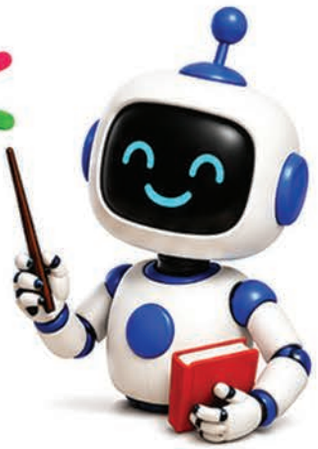
Metric	Calculation	Result
Euclidean	$\sqrt{[(7-3)^2 + (2-5)^2]} = \sqrt{[4^2 + (-3)^2]} = \sqrt{(16 + 9)} = \sqrt{25}$	**5.0**
Manhattan	$\|7-3\| + \|2-5\| = 4 + 3$	**7**

Where used:

- Maps: Euclidean (straight) vs Manhattan (streets)
- K-Nearest Neighbours uses Euclidean
- Word embedding similarity uses Cosine
- Image similarity uses Euclidean

"Manhattan" name comes from **NYC's grid streets!**

LINEAR REGRESSION



THE OLDEST ML ALGORITHM — STILL USED EVERYWHERE!

Goal: Find a LINE that best fits the data!

$$y = m \times x + c$$

where m = slope, c = y-intercept

Example: Predict house price from area

Area (sq ft)	Price (Lakhs)	Data point
500	30	(500, 30)
800	45	(800, 45)
1000	55	(1000, 55)
1500	80	(1500, 80)

Best fit: Price = 0.05 × Area + 5

For 1200 sq ft: Price = 0.05 × 1200 + 5 = **65 Lakhs**

What the model "learned":

- $m = 0.05$ → price grows ₹5000 per sq ft
- $c = 5$ → base price ₹5 Lakhs (land cost)

USES OF LINEAR REGRESSION

Sales forecasting, stock prediction, salary estimation, height/weight prediction, even crop yield from rainfall!