ML MODEL - TARGET/FEATURE - OVERFITTING / UNDERFITTING

ML Model

A model in machine learning is like a smart formula or math function that learns from data and can make predictions or decisions based on new input.

1800 Sq Ft ---> ? price

ML model – finds a relationship like (5000 per sq ft) and builds a pattern.

price = 0.05 * size
0.05 * 1800 = 90 L

Size	Price
1000	50 L
1500	75 L
2000	100 L

Target / Feature

Target— A target (also called as label or output) is the value that your model is trying to predict or learn.

Features— features are the input variables used to predict the target.
(size, location, bedrooms)

X - feature
y - target

Overfitting | Underfitting

Lazy but Confident

I’ll just read the last 1-page summary of the textbook. That should be enough for any question.

obsessed with details

I’ll memorize every single word from all previous exams — even the mistakes!

Balanced one

I’ll understand the concepts, do a few examples, and make sure I know how to solve new problems using logic.

🧠 In the exam, he couldn't answer most questions because they were not exactly like the summary.

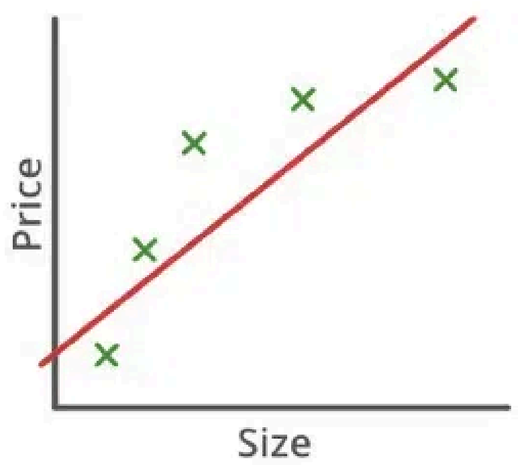
❌ Result: He failed.
🔎 Why? → His understanding was too simple → T**his is underfitting.**
📉 Error Type: High Bias — he assumed the exam would be too easy.

🧠 In the exam, the questions were slightly new, but her brain was filled with memorized patterns. She panicked.

❌ Result: She got confused and messed up.
🔎 Why? → She memorized the data instead of understanding it → **This is overfitting.**
📉 Error Type: High Variance — her model (mind) was too sensitive to the training data.

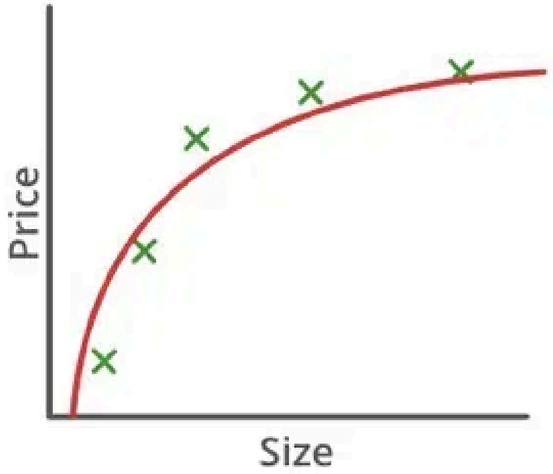
🧠 In the exam, even with new-style questions, Ravi was able to apply what he learned.

✅ Result: He passed with flying colors!
🔍 Why? → He balanced understanding with practice.
📉 Error Type: Low Bias + Low Variance = **Perfect generalization.**



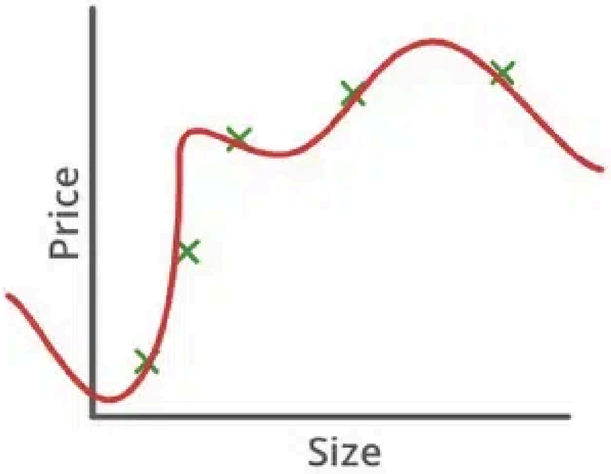
$\theta_0 + \theta_1 x$

High Bias
(Underfitting)



$\theta_0 + \theta_1 x + \theta_2 x^2$

Low Bias, Low Variance
(Goodfitting)



$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

High Variance
(Overfitting)

Today’s MCQ

Q1. Which of the following best describes an overfitted machine learning model?
A. It performs well on both training and test data
B. It is too simple to learn patterns
C. It performs well on training data but poorly on test data
D. It has low bias and low variance

Q2. A model that consistently predicts far from the actual value (even with different datasets) is said to have:
A. High Variance
B. Low Variance
C. High Bias
D. Low Bias

Q3. Vani is a student who memorizes every past question but struggles when questions change slightly. Her learning pattern is similar to:
A. Underfitting
B. High Bias
C. Overfitting
D. Low Variance

THANK YOU