

Lecture 5

Data in Each Learning Paradigm

Artificial Intelligence Programming
Department: Artificial Intelligence, College of Computer Science and Mathematics,
University of Mosul
Level: 2nd Year

Why Data Is the Foundation of AI

Every AI system is only as good as the data it learns from. A perfectly designed neural network trained on poor data will produce poor results. This principle is summarized as: *"Garbage In, Garbage Out."*

1.1 The Role of Data in Each Learning Paradigm

Supervised Learning

In supervised learning, the model learns $f(x) = y$ from labeled pairs. The quality and quantity of labels directly control what the model learns. Wrong labels create wrong decision boundaries, missing values cause biased estimates, and imbalanced classes cause the model to ignore minority classes entirely.

Aspect	What Good Data Provides	What Bad Data Causes
Labels	Accurate target for each sample	Misleads the model — learns wrong patterns
Volume	Generalizable learned function	Overfitting to small sample noise
Balance	Fair representation of all classes	Model ignores rare but important classes
Features	Relevant signal for prediction	Model learns noise as if it were signal

Unsupervised Learning

Unsupervised learning discovers hidden structure from data alone with no labels. Because there is no external validation signal, the algorithm is even more sensitive to data quality. Outliers distort cluster centroids in K-Means. Features on different scales cause distance-based algorithms to be dominated by high-magnitude features. Irrelevant features add noise to similarity measures, breaking natural groupings.

Reinforcement Learning

Reinforcement Learning (RL) is a type of machine learning where an **agent** learns how to behave in an **environment** by performing **actions** and receiving feedback in the form of **rewards or penalties**. The agent is not told what to do — it learns entirely by **trial and error**, gradually discovering which actions lead to the best outcomes.

This is fundamentally different from supervised learning (which needs labeled answers) and unsupervised learning (which finds structure with no feedback). In RL, the only feedback is a reward signal — a number that tells the agent whether its last action was good or bad.

The Core Learning Loop

```
Agent → observes State → chooses Action
Environment → changes to new State → gives Reward
Agent → learns from Reward → improves next Action
Repeat thousands of times until the agent masters the task
```

Key Components

Component	Role	Example — Robot Learning to Walk
Agent	The learner and decision maker	The robot itself
Environment	The world the agent interacts with	The floor, gravity, physical space
State	The current situation of the agent	Joint angles, balance, speed
Action	What the agent does at each step	Move left leg forward, bend knee
Reward	The feedback signal	+ 1 per step forward, – 5 for falling
Policy	The learned strategy	If balanced → step forward

Simple Analogy — Training a Dog

Think of training a dog. You do not explain rules in words. Instead: the dog sits when asked → reward (treat) → learns that sitting is good. The dog ignores the command → no reward → learns ignoring leads nowhere. Over many repetitions the dog masters the correct behavior. The RL agent learns *exactly* the same way — through repeated interaction and reward feedback.

❑ **Key Difference from Supervised Learning:** In supervised learning, a teacher provides the correct answer for every example. In RL, there is no teacher — only a reward signal that says 'good' or 'bad'. The agent must figure out WHY on its own.

How Data Quality Affects Reinforcement Learning

In RL, 'data' is the **stream of experiences** (state, action, reward) generated during the agent's interaction with the environment. Poor quality in this stream leads to wrong or failed learning.

Example — Teaching a Robot to Navigate a Warehouse

A company trains a robot to navigate a warehouse and deliver packages without hitting obstacles. Reward: +10 for successful delivery, -5 for hitting an obstacle, -1 per wasted second.

Case 1 — Good reward signal:

Situation	Reward	What Robot Learns
Delivers package correctly	+10	This path is correct — repeat it
Hits a wall	-5	This direction is dangerous — avoid it
Takes long unnecessary route	-1 per second	Find a shorter path
Smoothly avoids obstacle	+2	This avoidance behavior is good

□ **Result:** After thousands of interactions, the robot learns a safe and efficient policy — it delivers packages quickly and avoids all obstacles.

Case 2 — Faulty reward signal (bad data):

Sensors are broken: sometimes the robot gets +10 when it did NOT deliver (false positive), and -5 when it did nothing wrong (false negative).

Situation	Reward (Faulty)	What Robot Learns (Wrong!)
Hits obstacle	+10 (sensor error)	Hitting obstacles is good — do it more!
Correct delivery	-5 (sensor error)	Avoid making deliveries — they cause penalties!
Wanders randomly	+10 (false signal)	Random wandering is the best strategy!

□ **Result:** The robot learns a completely wrong policy and may deliberately crash into walls or avoid the task entirely. The model IS trained — but on corrupt data, so it learned the **WRONG** behavior.

Case 3 — Sparse reward signal:

1,000 actions taken → 999 receive reward = 0 → only 1 receives reward = +10

The agent has no idea which of the 999 actions contributed to success. Like sitting a 3-hour exam and only being told 'you passed' at the end with zero feedback on individual questions.

□ **Result:** Learning becomes extremely slow or fails completely. The robot wanders randomly for thousands of episodes before accidentally succeeding once.

Summary — Data Quality Effects in RL

Data Problem	Effect on Learning	Solution
Noisy / wrong reward signal	Agent learns incorrect behaviors	Fix sensors; use redundant reward sources
Sparse rewards	Learning extremely slow or fails	Reward shaping — give small intermediate rewards
Incomplete state (missing sensors)	Agent makes blind decisions	Add all relevant state variables and sensors
Only easy training scenarios	Agent fails on real complexity	Randomize training — vary obstacle layouts
Simulated vs real environment gap	Policy fails when deployed	Fine-tune with real-world experience

□ **The Golden Rule of RL Data:** In Reinforcement Learning, the reward signal IS the data. A wrong, noisy, or sparse reward is equivalent to corrupt labels in supervised learning. Designing a high-quality reward function is often the most critical step in any RL project.

1.2 The Data Pipeline — Overview

Every AI project follows a data pipeline before any model is trained:

1. Raw Data Collection — gather data from sensors, databases, APIs, or manual annotation
2. Data Inspection — understand distributions, types, missing rates, class balance
3. Data Cleaning — handle missing values, outliers, duplicates, inconsistencies
4. Data Transformation — normalize, encode categoricals, engineer features
5. Data Splitting — divide into training, validation, and test sets
6. Model Training & Evaluation — feed clean data to the algorithm

□ **Key Fact:** Studies consistently show that data scientists spend 60–80% of their time on data preparation (steps 1–5). Data quality is the dominant driver of model performance.

Data Problems: Types, Causes & Impacts

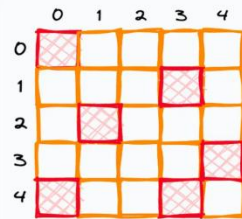
Before processing data, we must understand exactly what type of problem exists. Each problem type has different causes and requires different treatment methods.

2.1 Missing Values

Types of Missingness

Type	Definition	Example	Impact
MCAR — Missing Completely At Random	Missingness unrelated to any variable احتمال فقدان القيمة لا يعتمد على أي متغير في البيانات لا يعتمد على القيمة نفسها ولا على متغيرات أخرى. أي أن الفقدان مستقل تماماً عن البيانات.	Survey respondent skipped randomly انقطاع كهرباء أثناء إدخال البيانات. خطأ عشوائي في النظام. لا علاقة له بعمر المريض ولا بضغطه ولا بكونه مصاباً أو لا.	Least harmful — data is still representative أفضل سيناريو إحصائياً لأن حذف الصفوف لا يسبب تحيزاً كبيراً.
MAR — Missing At Random	Depends on other observed variables احتمال الفقدان يعتمد على متغيرات أخرى موجودة. لكن لا يعتمد على القيمة المفقودة نفسها بعد التحكم في المتغيرات الأخرى.	Men less likely to report weight المرضى الأكبر سناً أقل التزاماً بإجراء تحليل الكوليسترول. سبب الفقدان مرتبط بالعمر (Age) وليس مرتبطاً مباشرة بقيمة الكوليسترول نفسها.	Moderate — can be estimated from other variables إذا عرفنا العمر، يمكننا تقدير احتمال الفقدان.
MNAR — Missing Not At Random	Depends on the missing value itself مفقودة بسبب المتغير نفسه	High earners hide salary مرضى يعانون من ضغط دم مرتفع جداً يرفضون تسجيل قياساتهم لأنهم قلقون من النتائج.	Most harmful — creates systematic bias هو أخطر نوع لأن: لا يمكن استنتاج سبب الفقدان من البيانات المتاحة أي تعويض (Imputation) قد يكون منحازاً بشدة يؤدي إلى Bias خطير في التنبؤ الطبي

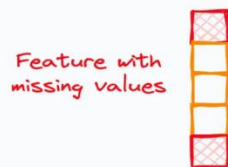
1) Missing completely at random (MCAR)



 blog.DailyDoseofDS.com

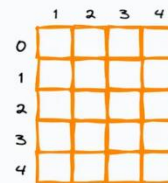
Data with randomly missing values

2) Missing at random (MAR)



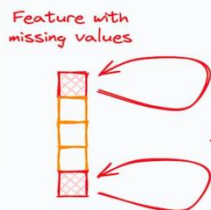
Feature with missing values

can be explained using



Observed data

3) Missing NOT at random (MNAR)



Feature with missing values

missingness is directly linked to missing value

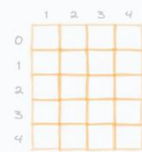
OR



Feature with missing values

missingness is linked to

Unobserved features



Mathematical Example — Missing Value Detection

Dataset of 8 house price records:

ID	Size (m²)	Rooms	Age (yrs)	Price (\$k)
1	120	3	10	250
2	NaN	4	5	310
3	95	NaN	20	180
4	200	5	3	NaN
5	140	3	NaN	270
6	NaN	NaN	15	220
7	110	2	8	195
8	175	4	12	NaN

Step 1 — Calculate missing rate per feature:

$$\text{Missing Rate} = (\text{Number of NaN values}) / (\text{Total rows}) \times 100\%$$

Feature	NaN Count	Missing Rate	Severity
Size (m ²)	2 out of 8	25%	Moderate
Rooms	2 out of 8	25%	Moderate
Age (yrs)	1 out of 8	12.5%	Low
Price (\$k) — TARGET	2 out of 8	25%	CRITICAL — must remove rows 4 and 8

❏ **Critical Rule:** If the TARGET variable (Price) has missing values, those rows **MUST** be removed. You cannot impute a label you are trying to predict. Remove rows 4 and 8 first, leaving 6 training samples.

2.2 Outliers

What Is an Outlier?

An outlier is a data point that lies far from the majority of the distribution. Outliers can be genuine rare events (valid and informative) or errors from data entry, sensor faults, or measurement mistakes. It is critical to determine which type before removing them.

هي قيمة تختلف بشكل كبير عن بقية البيانات بحيث لا تتبع النمط العام للتوزيع.