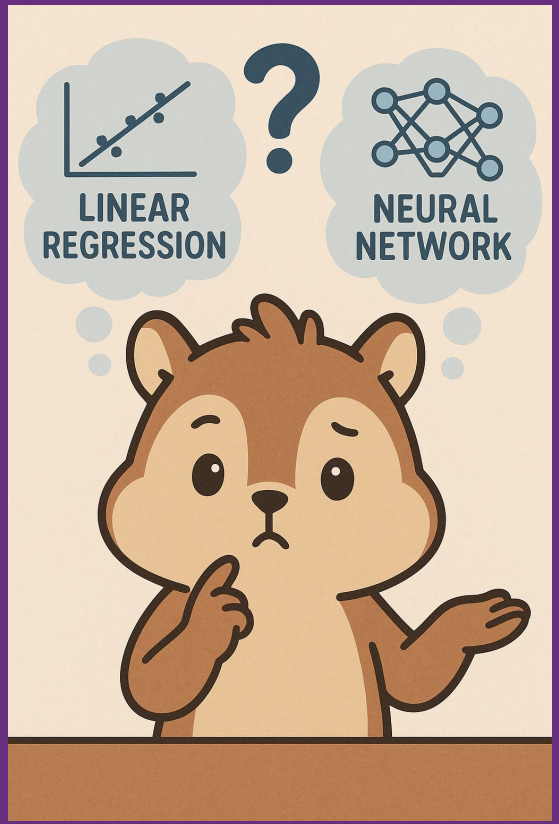


Image generated using Sora

Mila

10-06-2025

Choosing Appropriate ML Techniques

Prakhar Ganesh

About me

Prakhar Ganesh (he/him)

PhD student in Computer Science
at McGill University / Mila

Research in Fairness and Privacy in AI
& Multiplicity in AI

Goals today...

- Framing the Problem Statement
  - Supervised vs Unsupervised vs Reinforcement Learning
- Choosing the Model
  - Data Modality and Volume
  - Deployment considerations
- Evaluating your Solution

- Case Studies

Start with the Problem, Not the Solution!

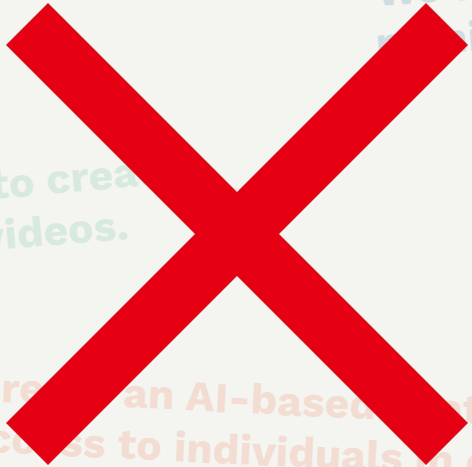
The Problem...

We will use image recognition to monitor customer satisfaction in physical stores.

We will use generative AI to create personalized workout videos.

We will create an AI-based chatbot to provide therapy access to individuals in our community.

The Problem...



We will use image recognition to
itor customer satisfaction in
physical stores.

We will use generative AI to crea
personalized workout videos.

We will cre an AI-based tbot to provide
therapy acc ss to individuals in our community.

The Problem...

“An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate problem.” - John Tukey

The Problem...

Not all problems need AI!

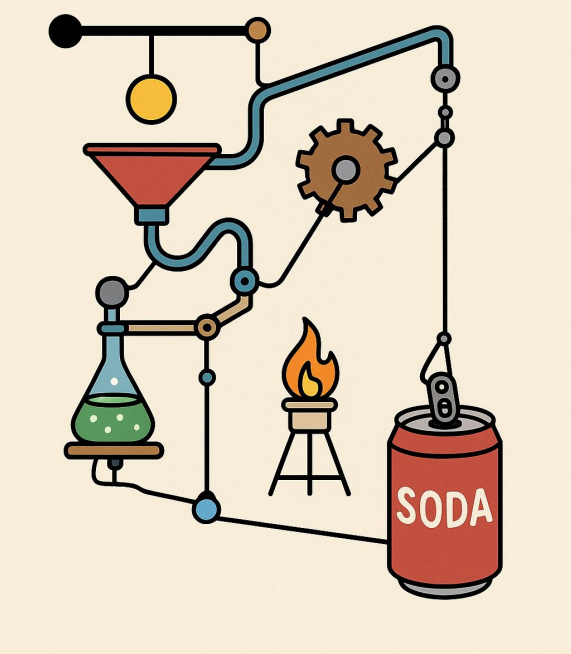


Image generated using Sora

The Problem...

Problem: ?

We will use image recognition to monitor customer satisfaction in physical stores.

The Problem...

Problem: ?

We will use image recognition to monitor customer satisfaction in physical stores.

Privacy Concerns!
Ambiguity of facial expressions
Camera positions
Overall Lack of Accuracy

The Problem...

Problem: Customers only shop at our store once and never return.

We will use image recognition to monitor customer satisfaction in physical stores.

Privacy Concerns!
Ambiguity of facial expressions
Camera positions
Overall Lack of Accuracy

The Problem...

Problem: Customers only shop at our store once and never return.

We will use image recognition to monitor customer satisfaction in physical stores.

Is it our prices?

Is it the experience?

...

Privacy Concerns!
Ambiguity of facial expressions
Camera positions
Overall Lack of Accuracy

The Problem...

Problem: Customers only shop at our store once and never return.

We will use image recognition to monitor customer satisfaction in physical stores.

Is it our prices?

Is it the experience?

...

Check prices of products against the market; see which products are being bought (No AI needed!!).

Privacy Concerns!
Ambiguity of facial expressions
Camera positions
Overall Lack of Accuracy

The Problem...

Problem: Customers only shop at our store once and never return.

We will use image recognition to monitor customer satisfaction in physical stores.

Privacy Concerns!
Ambiguity of facial expressions
Camera positions
Overall Lack of Accuracy

Is it our prices?

Is it the experience?

...

Use NLP to analyze customer reviews and find complaints.

Framing the Problem Statement

Supervised vs Unsupervised vs RL

Supervised vs Unsupervised vs RL

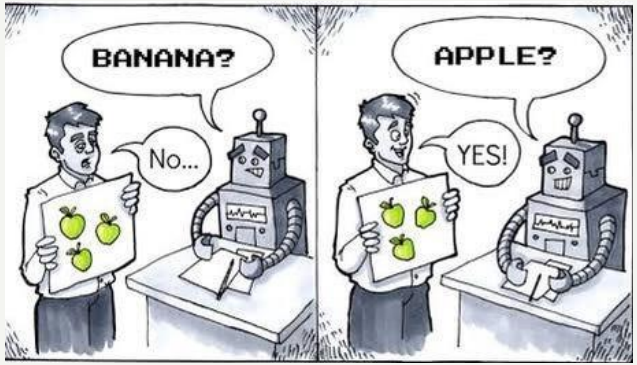
Do you have a clearly defined ‘target’ to predict?

Supervised vs Unsupervised vs RL

Do you have a clearly defined ‘target’ to predict?

Supervised Learning

Eg, Classification, Regression, Translation, etc.



Source: https://medium.com/datauniverse/unsupervised-learning-and-dimensional-reduction-663e00a3a086

Supervised vs Unsupervised vs RL

Do you have a clearly defined ‘target’ to predict?

Supervised Learning

Eg, Classification, Regression, Translation, etc.

Do you want to analyze data and extract patterns?

Supervised vs Unsupervised vs RL

Do you have a clearly defined 'target' to predict?

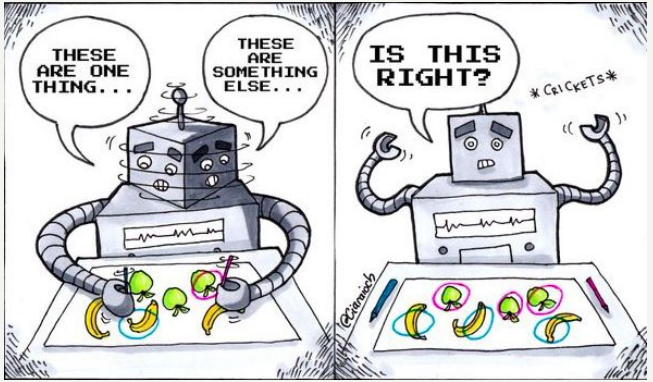
Supervised Learning

Eg, Classification, Regression, Translation, etc.

Do you want to analyze data and extract patterns?

Unsupervised Learning

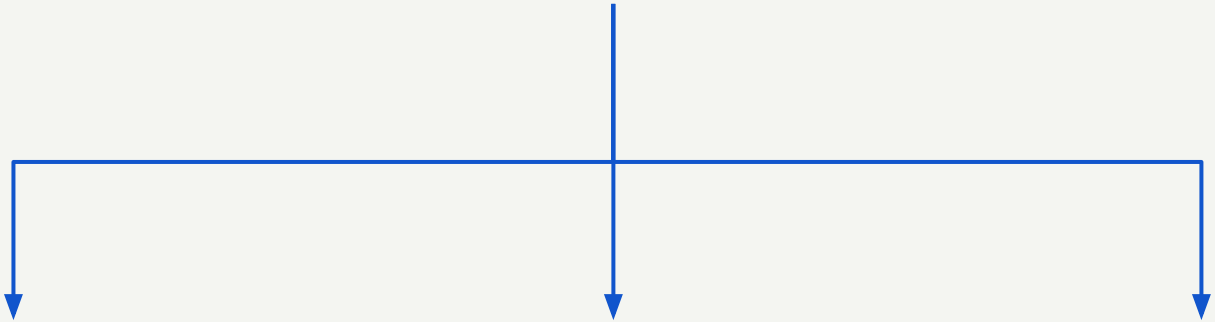
Eg, Clustering, Dimensionality reduction, etc.



Source: https://medium.com/datauniverse/unsupervised-learning-and-dimensional-reduction-663e00a3a086

Mila

Supervised vs Unsupervised vs RL



Do you have a clearly defined ‘target’ to predict?

Supervised Learning

Eg, Classification, Regression, Translation, etc.

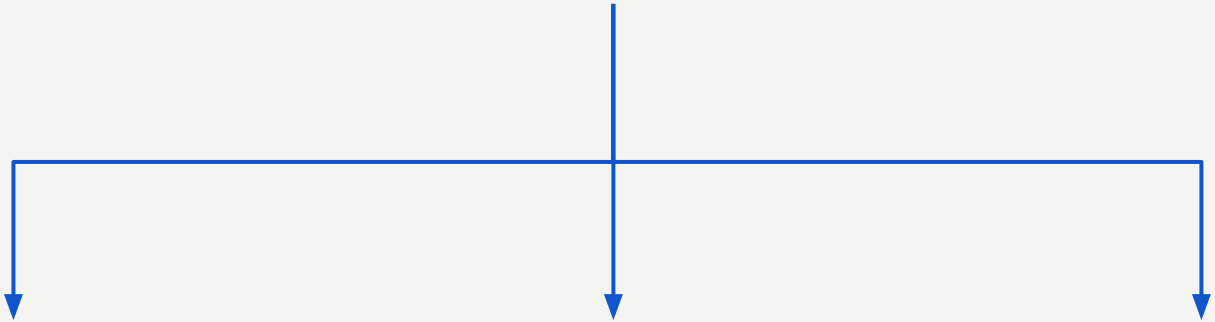
Do you want to analyze data and extract patterns?

Unsupervised Learning

Eg, Clustering, Dimensionality reduction, etc.

Do you have an overall goal in an active environment?

Supervised vs Unsupervised vs RL



Do you have a clearly defined ‘target’ to predict?

Supervised Learning

Eg, Classification, Regression, Translation, etc.

Do you want to analyze data and extract patterns?

Unsupervised Learning

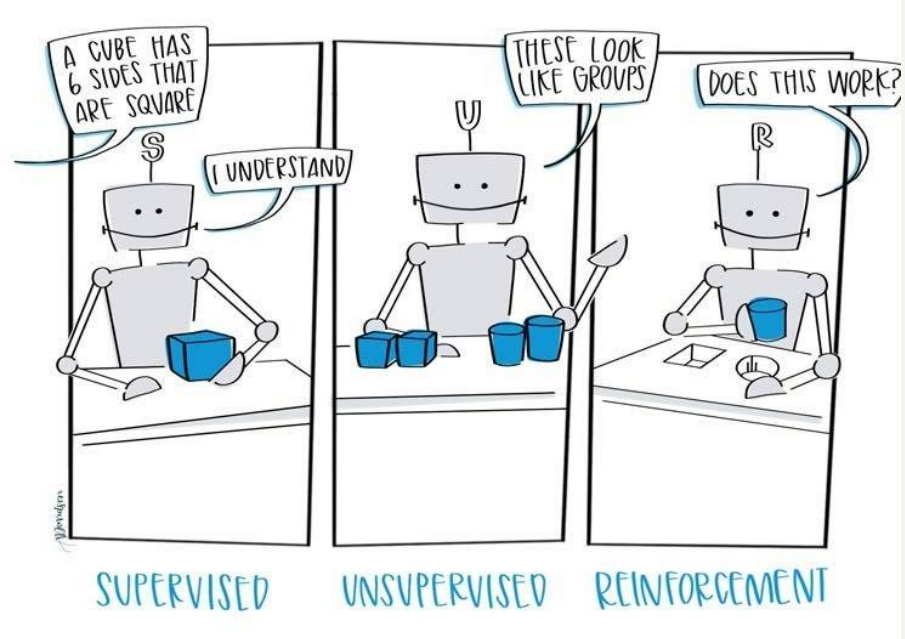
Eg, Clustering, Dimensionality reduction, etc.

Do you have an overall goal in an active environment?

Reinforcement Learning

Eg, Robotics, Game playing agents, etc.

Supervised vs Unsupervised vs RL



Source:
https://kasunprageethdissanayake.medium.com/artificial-intelligence-2-supervised-learning-unsupervised-learning-and-reinforcement-learning-7bf00c732e99

Mila

Examples

Examples

Learn how to recognize the species of a bird from its photo

Examples

Learn how to recognize the species of a bird from its photo

Supervised

Examples

Learn how to recognize the species of a bird from its photo — Supervised

Content recommendation on Youtube

Examples

Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL

Examples

Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL
Analyzing shopping behaviour to find which items are usually bought together	

Examples

Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL
Analyzing shopping behaviour to find which items are usually bought together	Unsupervised

Examples

Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL
Analyzing shopping behaviour to find which items are usually bought together	Unsupervised
Self-driving cars	

Examples

Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL
Analyzing shopping behaviour to find which items are usually bought together	Unsupervised
Self-driving cars	RL

Examples

Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL
Analyzing shopping behaviour to find which items are usually bought together	Unsupervised
Self-driving cars	RL
Finding anomaly in banking behaviour to detect fraud	

Examples

Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL
Analyzing shopping behaviour to find which items are usually bought together	Unsupervised
Self-driving cars	RL
Finding anomaly in banking behaviour to detect fraud	Unsupervised

Examples

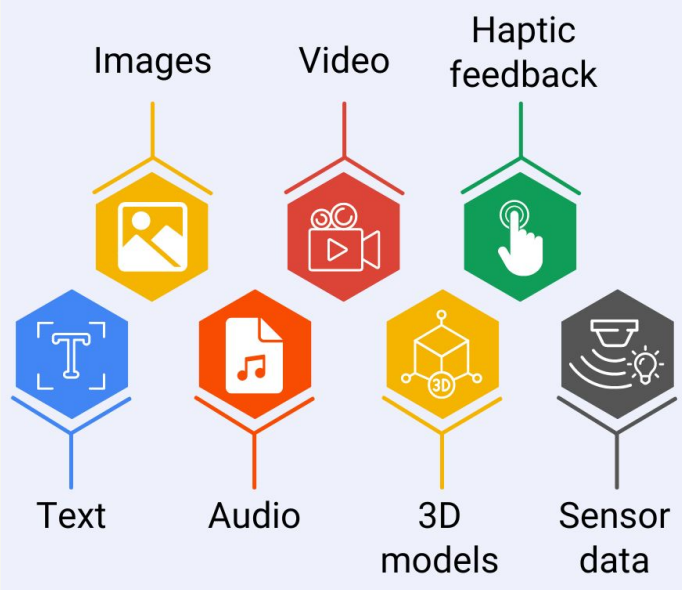
Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL
Analyzing shopping behaviour to find which items are usually bought together	Unsupervised
Self-driving cars	RL
Finding anomaly in banking behaviour to detect fraud	Unsupervised
Identify proper nouns and ‘personal identifiers’ to avoid privacy concerns	

Examples

Learn how to recognize the species of a bird from its photo	Supervised
Content recommendation on Youtube	RL
Analyzing shopping behaviour to find which items are usually bought together	Unsupervised
Self-driving cars	RL
Finding anomaly in banking behaviour to detect fraud	Unsupervised
Identify proper nouns and ‘personal identifiers’ to avoid privacy concerns	Supervised

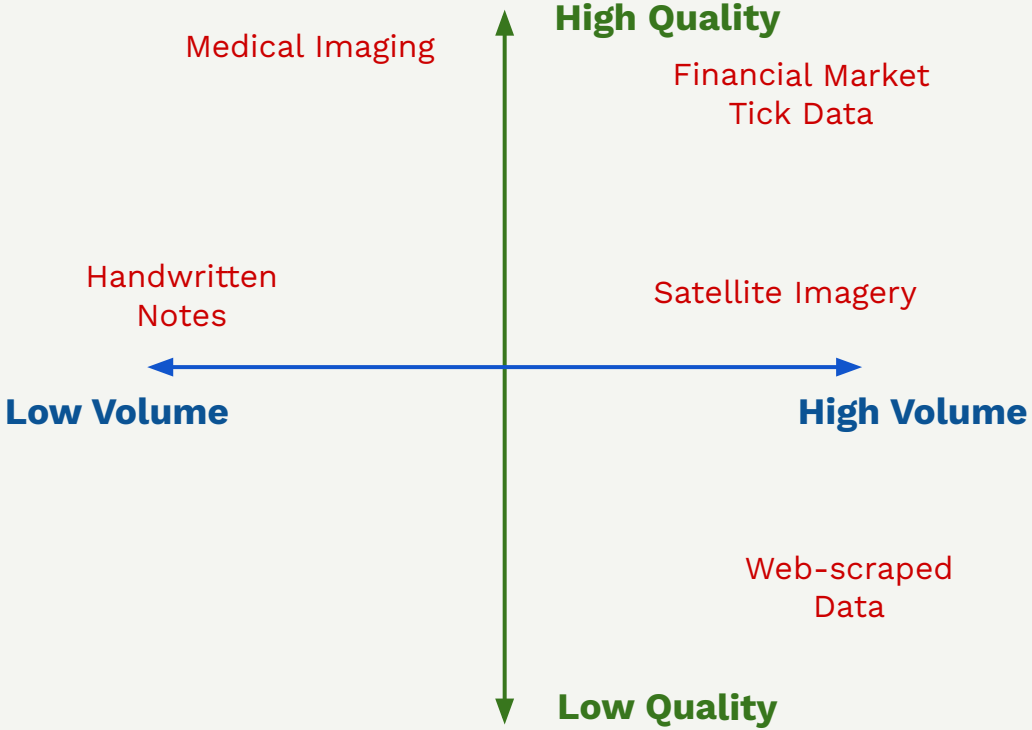
Choosing the Model

Data “Modality”



Source: https://datasciencedojo.com/blog/multimodality-in-llms/

Data Volume and Quality



Deployment Constraints

Integration Constraints

Interpretability

Latency and Network Requirements

Storage Constraints

Data Privacy and Security Concerns

Training Feasibility

Evaluating your Solution

Choosing Evaluation Metric

What matters to the stakeholders?

Data Splits and Overfitting

Imagine you are studying for a course and preparing for the exam.

Data Splits and Overfitting

Imagine you are studying for a course and preparing for the exam.

Course objective

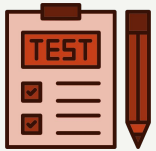
Data Splits and Overfitting

Imagine you are studying for a course and preparing for the exam.

Course objective

Your notes

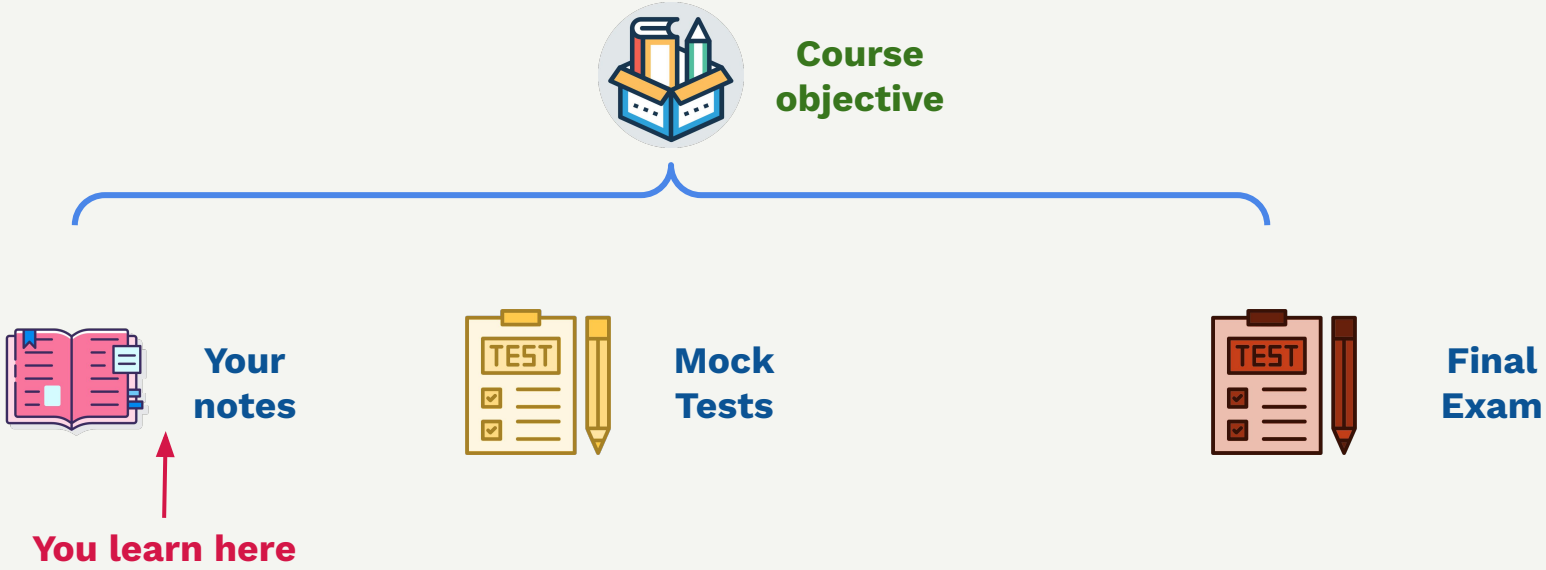
Mock Tests



Final Exam

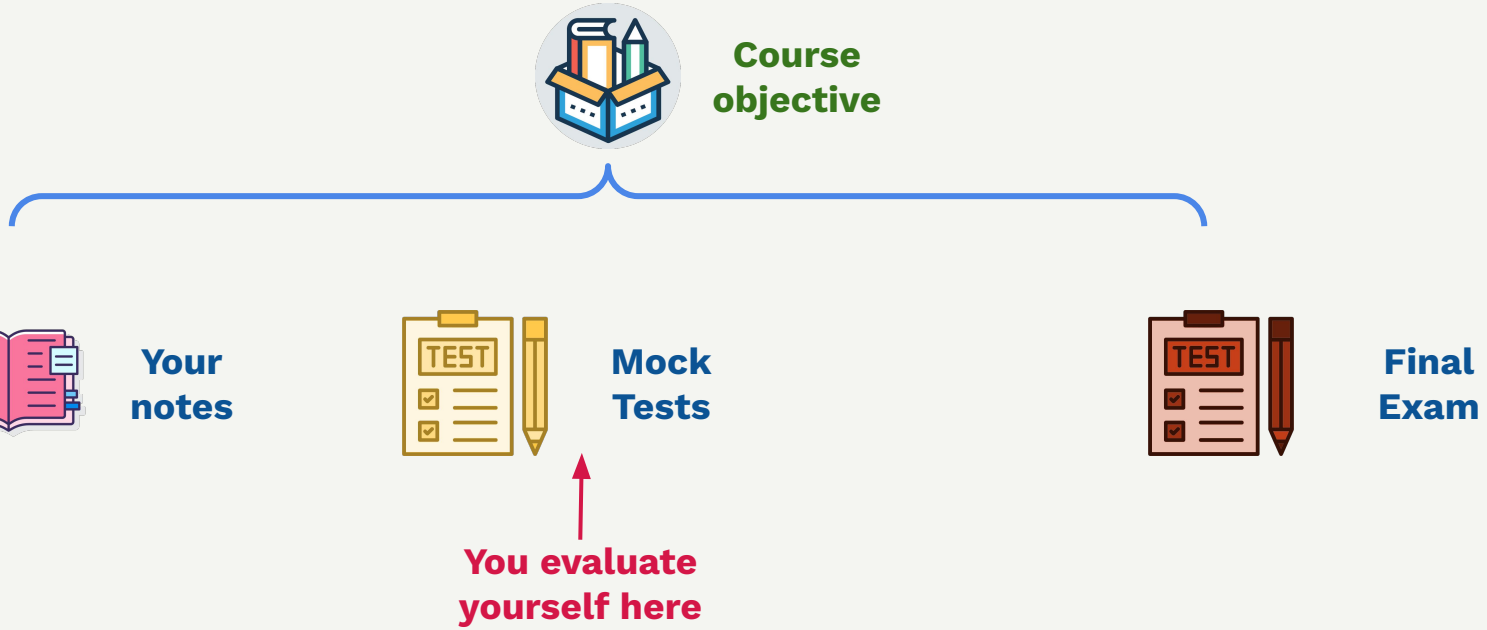
Data Splits and Overfitting

Imagine you are studying for a course and preparing for the exam.



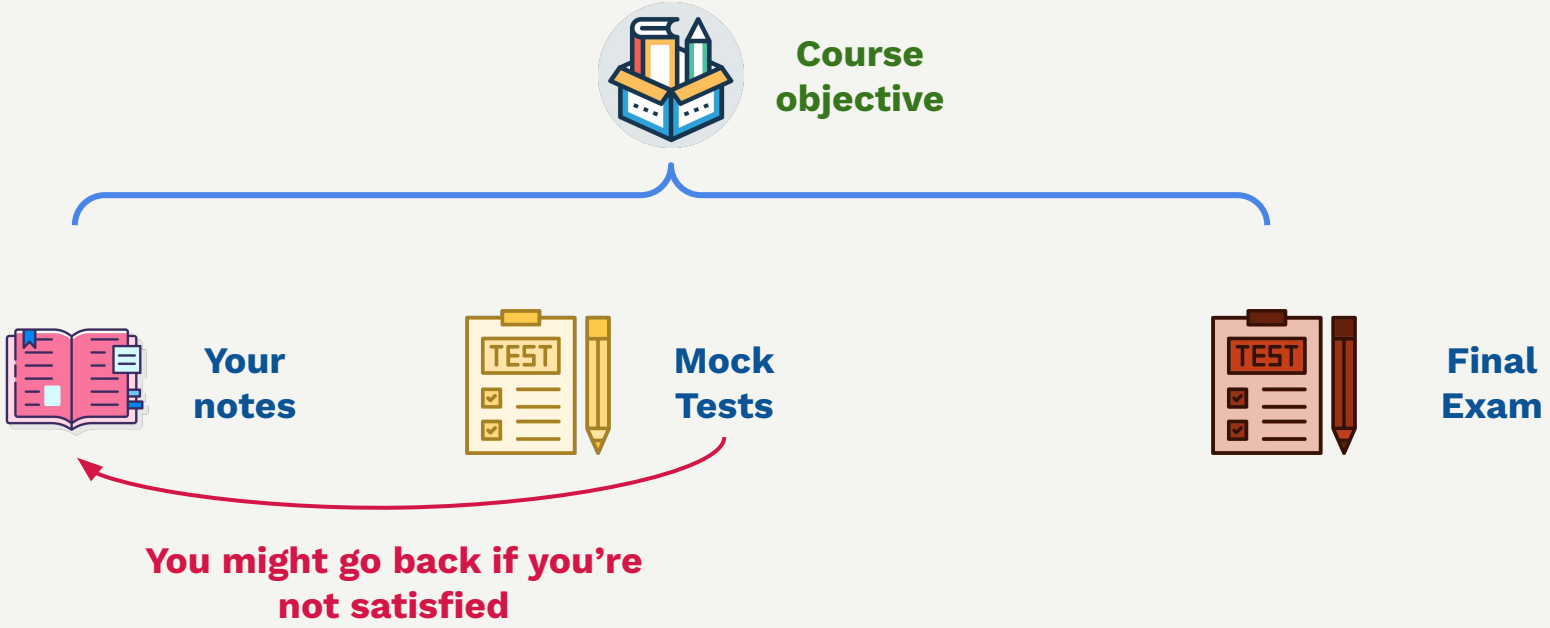
Data Splits and Overfitting

Imagine you are studying for a course and preparing for the exam.



Data Splits and Overfitting

Imagine you are studying for a course and preparing for the exam.



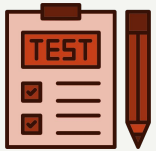
Data Splits and Overfitting

Imagine you are studying for a course and preparing for the exam.

Course objective

Your notes

Mock Tests



Final Exam

Final grades here

Data Splits and Overfitting

Imagine you are studying for a course and preparing for the exam.

Course objective

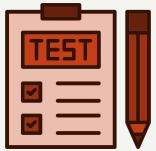
Data Distribution

Your notes

Training Data

Mock Tests

Validation Data



Final Exam

Test Data

Case Studies

CS1: Reducing Customer Churn

A subscription-based fitness app is experiencing a gradual drop in its monthly active users and subscription renewals.

CS2: Organizing Company Documents

A company has a growing internal knowledge base with documents across multiple departments. Employees complain that it's hard to find relevant information quickly.

CS3: Pricing for a Ride-Sharing App

Riders are increasingly complaining about sudden price surges and unpredictable ride fares during peak hours on the ride-sharing platform. At the same time, driver availability fluctuates, leading to inconsistent service quality and lost revenue opportunities.