



THE LEARNING ROADMAP

Machine Learning in the Age of AI

Equations, data, learning algorithms and intelligent systems—connected through a small experiment you can reproduce.



Equation → Data → Train → Check → Predict

What you should be able to explain

- How a mathematical model becomes a fitted predictor.
- Why EDA and data timing come before algorithm choice.
- How KNN, SVM, XGBoost, K-means and LSTM differ.
- How to interpret scores, check generated claims and reason about agents.

Opening film: BabyX First Words — University of Auckland Laboratory for Animate Technologies. An expressive interface prompts questions; evidence establishes capability.



01 / THE MODEL

A differential equation with an unknown rate

Liquid begins at 80°C in a 25°C room. Assume the room temperature and effective vessel properties remain constant. Temperature changes in proportion to its difference from the room.

$$dT/dt = -k(T - 25), \quad T(0) = 80$$

$$T(t) = 25 + 55 \exp(-kt)$$

t is in minutes; T is in °C; k is in min^{-1} . The structure comes from the physical model. The unknown k comes from fitting observations.

Time (min)	Observed °C	Data role
0	80	Training
5	70.2	Training
10	62.4	Training
15	54.9	Training
20	50.1	Training
25	44.9	Training
30	41.8	Training

All observations are synthetic. The 35-minute value (38.4°C) and 40-minute value (36.2°C) are held out for a later check. They do not determine k .

A low error cannot establish that a different cup, liquid or room will follow the same equation.



02 / THE EXPERIMENT

Fitting, training and a new prediction

1. **Initial guess:** $k = 0.015 \text{ min}^{-1}$. Calculate predicted temperatures at the training times.
2. **Loss:** mean squared error = average of $(\text{prediction} - \text{observation})^2$.
3. **Training:** repeatedly update k to reduce this loss. The notebook optimises $\log(k)$, so k remains positive.
4. **Check:** compare the fixed fitted model with the two held-out observations.
5. **Inference:** put a new time into the fitted equation without changing k .

Quantity	Calculated result
Initial training MSE	175.461 °C ²
Fitted k after 180 updates	0.03978592 min ⁻¹
Final training MSE	0.09379 °C ²
Held-out MAE (two observations)	0.133°C

A new input: 45 minutes

$$T(45) = 25 + 55 \exp(-0.03978592 \times 45)$$

$$\approx 34.18^\circ\text{C}$$

Changing k is additional training. Changing t while keeping k fixed is inference. This estimate extrapolates beyond the 0–30 minute training interval.

Run the supplied cooling notebook or use the interactive experiment. Training uses learning rate 0.001 and 180 gradient updates.



03 / CHOOSING THE METHOD

Data, tasks and model families

The separate campus-energy CSV has 600 fictional independent scenarios. Temperature and expected occupancy each have 24 missing entries. EDA checks meanings, units, ranges, missingness, duplicates, relationships and availability at prediction time.

Method	Idea	Important distinction
Linear / logistic	Weighted input relationships	Number / class probability
KNN	Nearby labelled cases	Scaling and distance matter
SVM	Separating margin, optional kernel	Supervised method
Trees / XGBoost	Splits / boosted trees	Often useful for tabular data
K-means	K groups without outcome labels	Clusters need interpretation
LSTM	Gated state through ordered steps	Sequence neural architecture
Transformer	Attention between representations	Architecture used in foundation models

These are different tasks, methods and architectures. Compare models only after making inputs, timing and evaluation conditions comparable.

Campus validation: mean baseline 19.53 kWh MAE; permitted-input linear model 7.46 kWh. A deliberately leaked end-of-day meter gives 0.0 because it copies the target too late. A perfect number can be invalid.

Use later-period tests for a real time series. The campus teaching rows are independent fictional scenarios.



04 / INTELLIGENT SYSTEMS

Foundation models, scores and agents

GPT and Gemini are model families. As of 17 October 2026, official catalogues list GPT-6 Astra and Gemini 4 Argon. A typical development story includes pretraining, further tuning, evaluation and deployment; exact proprietary recipes may be only partly public.

What does a score mean?

Ask which benchmark, dataset, metric, denominator, date and conditions produced the result. A high score does not certify a different local task. Accuracy, MAE, pass rate and human preference measure different things.

Hallucinations

A fluent answer can contain an unsupported claim, invented citation or false number. Open the source and verify the exact claim. Use deterministic calculations for arithmetic. Retrieval helps but does not guarantee truth.

Agentic systems

Goal → **Plan** → **Tool** → **Observe** → **Revise**

A complete agent includes tools, state, permission boundaries, stop rules and review. A safe teaching example reads a synthetic CSV, counts missing entries and drafts a report for a person to approve.

Engineering work

AI can assist with bounded tasks. Engineers still frame the problem, test failures, secure integrations and take responsibility for deployment. Avoid assuming that a capability demonstration predicts job replacement.

Primary references: OpenAI models · Google DeepMind model cards · Attention Is All You Need



05 / WORKSHEET AND RESOURCES

Your next experiment

Choose a useful ML application. Write a small plan that another person can test.

Decision and prediction time

One likely failure

Target and two usable inputs

What evidence is missing?

Baseline and metric

Who reviews the output?

Questions for discussion

1. Why is fitting k training, while substituting a time is inference?
2. Why can a lower training error still fail in a different room?
3. What separates K-means from KNN?
4. Why might a simple baseline beat an LSTM on a small task?
5. How would you check a generated citation?

Workshop 01 — notebooks, dataset and model-card exercise

MLOps Academy — versioning, evaluation and deployment

MindForgeAI — 13-project catalogue

Akash Chatake — portfolio

Method references: scikit-learn · XGBoost · Keras LSTM. All numerical cooling and campus examples in this guide are synthetic.