SH Dr. Soufiane Hourri

drsoufiane.com

DUT · SEMESTER 3 · COMPUTER SCIENCE

Machine Learning

Practice by hand · Linear regression (Session 03)

Dr. Soufiane Hourri, PhD, HDR

Associate Professor & AI Consultant

École Supérieure de Technologie de Safi · Cadi Ayyad University

2026-2027

PRACTICE · OVERVIEW

One hour of practice, by hand

01 Least squares by hand · 13 min

02 From r to the line · 8 min

03 Comparing models on a test set · 13 min

04 Reading a multiple regression · 13 min

05 Checking a model · 10 min

HOW WE WORK

In pairs, with paper and a calculator.

Every result with its unit and one sentence of interpretation.

Formula card

Quantity	Formula
Line	$\hat{y} = b_0 + b_1 x$
Slope	$b_1 = \Sigma (x_i - \bar{x})(y_i - \bar{y}) / \Sigma (x_i - \bar{x})^2$ or $b_1 = r \times s(y) / s(x)$
Intercept	$b_0 = \bar{y} - b_1 \times \bar{x}$
Residual	$e_i = y_i - \hat{y}_i$ (the residuals of a least-squares line add up to 0)
MAE · RMSE	$MAE = \Sigma \lvert e_i \rvert / n$ $RMSE = \sqrt{(\Sigma e_i^2 / n)}$
R^2	$R^2 = 1 - \Sigma e_i^2 / \Sigma (y_i - \bar{y})^2$ (simple regression: $R^2 = r^2$)

Multiple regression: $\hat{y} = b_0 + b_1x_1 + \ldots + b_px_p$. Each b_j is read "all else being equal", with its unit.

Exercise 1 · Least squares by hand

DATA (EXAMPLE)

(Experience in years, salary in thousand MAD): (1, 3), (2, 5), (3, 4), (4, 7), (5, 9)

TIME

13 min

QUESTIONS

1. Compute $\bar{x}$, $\bar{y}$, the slope b_1 and the intercept b_0
2. Interpret b_1 and b_0, with units
3. Residuals, then MAE, RMSE and R^2
4. Predict for 6 years. And for 30 years?

TIP

Table: $x_i - \bar{x}$, $y_i - \bar{y}$, their product, $(x_i - \bar{x})^2$.

Exercise 2 · From r to the line

DATA (EXAMPLE)

Study hours: mean 6 h, s = 2 h. Grade: mean 12/20, s = 3. Correlation r = 0.8.

TIME

8 min

QUESTIONS

1. Compute b_1 and b_0 of the regression line
2. Predict the grade for 8 h, then for 6 h. Comment
3. Compute R^2 and interpret it
4. What would the line be if r = 0?
5. And if r = −0.8?

ANSWER FORMAT

Each result with its unit; one sentence for questions 3 to 5.

Exercise 3 · Predictions on a test set

Test row	1	2	3	4	5
True value y	10	12	15	18	20
Model A	11	12	14	17	21
Model B	10	12	15	18	25
Baseline	14	14	14	14	14

Example data. Baseline = mean of the training set (14) predicted for every row.

EXERCISE 3 · SESSION 03

Exercise 3 · Which model is better?

GOAL

Compute the three metrics for each model and choose.

TIME

13 min

QUESTIONS

1. Errors $e_i = y_i - \hat{y}_i$ for A, B and the baseline
2. MAE, RMSE and R^2 for each
3. Which model do you choose? Why?
4. Why is B's RMSE so much larger than its MAE?
5. The baseline has $R^2 < 0$. How is that possible?

TIP

Compute $\Sigma (y_i - \bar{y})^2$ once: it is the same for all three.

Exercise 4 · Reading a multiple regression

MODEL (EXAMPLE)

price = 150 + 9 × area + 120 × Marrakech − 40 × Agadir

TIME

13 min

QUESTIONS

1. Predict an 80 m² apartment in Safi, Marrakech, Agadir
2. Interpret the three coefficients, all else equal
3. 100 m² in Safi or 80 m² in Marrakech: which costs more?
4. Why has Safi no coefficient?
5. Can the model price a 400 m² villa?
6. Area in cm²: what becomes the coefficient 9?

UNITS

Price in thousand MAD, area in m² (training: 40 to 150). Reference city: Safi. A dummy is 1 for its city, 0 otherwise.

EXERCISE 5 · SESSION 03

Exercise 5 · Checking a model

DATA (EXAMPLE)

Residuals of a line, for x = 1 to 6: +2, 0, −1, −1, 0, +2

TIME

10 min

QUESTIONS

1. Sketch the residuals against x. What pattern? What does it mean?
2. What would residuals growing with x mean?
3. R^2 (train, test) of three models: (0.62, 0.60), (0.81, 0.78), (0.97, 0.41). Which overfits? Which do you choose?

ANSWER FORMAT

A small sketch and one sentence per question.

To finish: one question

QUESTION

In salary-hat = 1.4 + 1.4 × years, what does the slope 1.4 mean?

A

Every employee earns exactly 1,400 MAD more each year

B

One more year of experience causes a raise of 1,400 MAD

C

One more year is associated with 1,400 MAD more predicted salary

D

1.4% of the predictions are exact

SH

Well done · Questions?

Keep your sheets: these calculations come back in the exam.

Dr. Soufiane Hourri, PhD, HDR

drsoufiane.com