Understanding Artificial Intelligence from Scratch

Solutions manual and marking guidance (for learners checking their work, and for instructors).

O1: Explain in plain words what artificial intelligence and machine learning are.

Q01 [Foundational · short answer] In plain words, what does an AI system do according to the OECD definition used in the lesson?

Answer: It works out, from its input, how to produce outputs such as predictions, content, recommendations or decisions.

1. Key words: input, works out (infers), output.

Q02 [Foundational · true/false] True or false: all AI systems are robots.

Answer: False

1. Most AI is software with no physical body.

Q03 [Foundational · multiple choice] Machine learning is best described as: (a) writing every rule by hand (b) finding patterns in examples (c) building robots (d) storing files

Answer: (b)

1. Learning from labelled examples instead of hand-written rules.

Q04 [Core · conceptual] Why did rule-based spam filters struggle?

Answer: Scammers changed their wording and genuine messages contained 'spam words', so fixed rules could not keep up.

1. Learning from many examples adapts better.

Q05 [Core · classification] Rule-based or learning? (a) a calculator adding numbers (b) a phone recognising your face (c) a tax form applying fixed rates

Answer: (a) rules, (b) learning, (c) rules

1. Exact, simple logic suits rules; complex patterns suit learning.

Q06 [Applied · explanation] Give one task better solved with rules and one better solved with machine learning, and justify each.

Answer: Example: VAT calculation (rules: exact, simple); recognising handwritten digits (learning: too many variations to list).

1. Judge by complexity of the pattern and need for exactness.

O2: Describe how training data turns into a model that makes predictions.

Q07 [Foundational · short answer] What is a label in machine learning?

Answer: The correct answer attached to a training example.

1. Example: 'apple' attached to a fruit's measurements.

Q08 [Foundational · multiple choice] What do we call using a trained model to make a prediction on new data? (a) training (b) labelling (c) inference (d) overfitting

Answer: (c)

1. Training builds the model; inference uses it.

Q09 [Core · numerical] Using the rule 'apple if redness >= 0.1 x weight - 13', classify a fruit of 210 g with redness 5.

Answer: orange

1. Threshold 0.1 x 210 - 13 = 8; redness 5 < 8, so orange.

Q10 [Core · numerical] A model makes 50 predictions on test data and 42 are correct. What is its accuracy?

Answer: 84%

1. 42 / 50 = 0.84 = 84%.

Q11 [Applied · conceptual] Explain overfitting using the example of a student who memorises past exam answers.

Answer: The student scores well on familiar questions but fails new ones; a model that memorises training data does the same on new data.

1. Test data reveals overfitting.

Q12 [Applied · analysis] Of 1,000 emails, 950 are genuine. A model labels every email 'genuine'. Calculate its accuracy and explain why it is still useless.

Answer: 95% accurate, but it catches no spam at all.

1. 950 / 1000 = 95%. Accuracy hides that it never detects the class we care about.

Q13 [Advanced · design] You want to build a model that predicts whether a plant needs water. Name three features, the label, and how you would split the data for testing.

Answer: Example features: soil moisture, temperature, days since watering; label: needs water yes/no; keep e.g. 20% of examples aside as test data.

1. Features must be measurable before the decision.

O3: Understand, at a simple level, how generative AI creates text.

Q14 [Foundational · short answer] What simple task is a large language model trained to do?

Answer: Predict the next token (word or part of a word).

1. Repeated many times, this produces whole texts.

Q15 [Foundational · true/false] True or false: a language model looks up every fact in a trusted database before answering.

Answer: False

1. It predicts likely words; it does not automatically check facts.

Q16 [Core · numerical] Next-word probabilities are: tea 0.5, coffee 0.3, juice 0.15, soup x. Find x.

Answer: 0.05

1. Probabilities add to 1: 1 - (0.5 + 0.3 + 0.15) = 0.05.

Q17 [Core · multiple choice] What is a 'hallucination' in generative AI? (a) a picture of a dream (b) a confident but false output (c) a type of neural network (d) a training dataset

Answer: (b)

1. It sounds plausible but is wrong.

Q18 [Applied · explanation] Explain, step by step, how a chatbot produces a sentence of reply.

Answer: It predicts probabilities for the next token, chooses one, appends it, and repeats with the longer text until it stops.

1. Each step depends on everything written so far.

Q19 [Advanced · analysis] Why can a language model write fluent text about a topic yet still make factual mistakes about it?

Answer: Fluency comes from learning patterns of language; truth requires checking facts, which next-token prediction alone does not do.

1. Hence the need to verify important claims.

O4: Recognise what AI does well, where it fails, and how to use it responsibly.

Q20 [Foundational · short answer] Name two things AI does well.

Answer: Examples: recognising patterns in large data sets; translation; recommending; flagging unusual payments; assisting with medical scan review.

1. Tasks with many examples and clear patterns.

Q21 [Foundational · multiple choice] Which should you NOT paste into a public AI tool? (a) a recipe question (b) your bank password (c) a request to explain photosynthesis (d) a poem prompt

Answer: (b)

1. Never share sensitive or private information.

Q22 [Core · explanation] How can bias get into an AI system?

Answer: Through training data that is unbalanced or reflects past prejudice; the model learns and repeats those patterns.

1. Better, more representative data and testing across groups help.

Q23 [Core · true/false] True or false: AI tools used in hospitals should make final diagnoses without a doctor.

Answer: False

1. A qualified human keeps responsibility for important decisions.

Q24 [Applied · application] A chatbot gives you a statistic for an essay. Describe how you would use it responsibly.

Answer: Check it against a reliable original source, cite that source, and follow the rules on declaring AI use.

1. Treat AI as an assistant, not an authority.

Q25 [Advanced · evaluation] A company plans to use AI to screen job applications automatically. List three risks and one safeguard for each.

Answer: Bias (audit outcomes by group); errors on unusual CVs (human review of rejections); lack of transparency (explain the process to applicants).

1. Keep a human accountable for decisions about people.

Marking scheme: Assignment 1: AI in my day (individual, about 500 words)

- Correct identification of AI systems: 2 marks
- Input-model-output explained for three systems: 6 marks
- Sensible testing plan: 3 marks
- Risks and safeguards realistic and specific: 6 marks
- Clarity: 3 marks (total 20)

Marking scheme: Assignment 2: check the chatbot (pairs)

- Questions and sources recorded: 3 marks
- Accurate checking: 4 marks
- Accuracy calculated correctly: 2 marks
- Explanation of errors: 3 marks (total 12)