Artificial intelligence, machine learning and robotics

5.2.2 · Issues and impact · revision notes

Artificial intelligence (AI), machine learning and robots are making more decisions and doing more jobs. This subtopic covers the ethical and legal issues they raise: accountability, safety, algorithmic bias and legal liability. Questions on Paper 1 are usually set in a context, from 2-mark 'Explain' questions to 6-mark 'Discuss' questions.

Key facts

- **AI**: computer systems that do tasks that normally need human intelligence.
- **Robot**: a computer-controlled machine that carries out physical tasks automatically.
- **Algorithmic bias**: consistently unfair results for some groups, usually caused by unrepresentative or prejudiced training data.
- **Legal liability**: who is legally responsible (e.g. pays compensation) when an AI or robot causes harm.
- Safeguards: representative training data, testing for bias, safety features, a human in the loop.
- **Machine learning**: AI that learns patterns from training data instead of following hand-written rules.
- **Autonomous**: able to make decisions and act without human control.
- **Accountability**: who must answer for a decision made by an AI or robot.
- **Safety**: autonomous machines can cause physical harm if they fail, meet something new or are hacked.

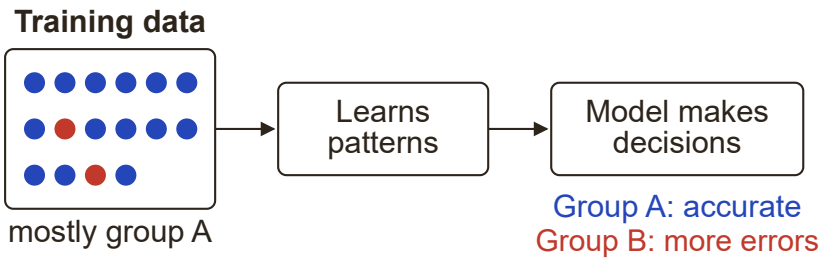
Notes

Key terms

- **Artificial intelligence (AI)**: computer systems that carry out tasks that normally need human intelligence, such as recognising faces or speech, translating or making decisions.
- **Machine learning (ML)**: a type of AI that learns patterns from large amounts of **training data**, instead of being programmed with rules for every situation. It usually improves as it is given more data.
- **Robot**: a computer-controlled machine that carries out physical tasks automatically. Most robots use sensors; some use AI to act without human control (they are **autonomous**).
- Benefits: AI works quickly, all day and consistently; robots do dangerous, repetitive or precise jobs, such as bomb disposal. Both can also replace some people's jobs.

Algorithmic bias

- An ML system can only learn from its training data. If the data is unrepresentative (e.g. mostly one group of people) or reflects past human prejudice, the system learns the same unfairness.



A model learns only from its training data. Blue dots are group A, red dots group B: too few red examples leads to biased results.

- **Algorithmic bias** is when a system consistently produces unfair results for some groups, e.g. face recognition that misidentifies some ethnic groups more often, or a recruitment tool that favours one gender.
- It is hard to spot, because people tend to trust a computer's decision and the system may not be able to explain it.
- To reduce it: use large, representative training data, test the results for different groups and have humans review important decisions.

Safety

- Autonomous machines act in the real world, so a wrong decision can injure or kill, e.g. a self-driving car failing to detect a cyclist.
- An AI may behave unpredictably when it meets a situation that was not in its training data.
- Machines must be thoroughly tested, have safety features such as emergency stops and be protected from hacking.

Accountability and legal liability

- **Accountability**: who is responsible and must answer for a decision made by an AI or robot.
- **Legal liability**: who is legally responsible, e.g. who must pay compensation, if an AI or robot causes harm or breaks the law.
- This is hard to decide: the programmers, the manufacturer, the owner and the user could all be involved, and the machine made the decision itself.
- With machine learning, nobody wrote the rule that caused the decision, and the system may not be able to explain why it acted as it did (a 'black box'). Most laws were written for human decision-makers.
  Grade 7+ only
- Keeping a human involved in important decisions (medical, jobs, loans, policing) means that a person stays accountable.

How to answer each type of question

Define or describe a key term

1 to 2 marks | Grade 4

1. Give the meaning in one clear sentence.
2. For 2 marks, add a second point, such as how it works or what it allows.
3. Avoid vague answers such as 'a computer that thinks for itself'.

Example. Describe what is meant by 'machine learning'. (2 marks)

Model answer. A type of AI in which a computer system learns patterns from training data (1), so that it can make predictions or decisions without being given rules for every situation (1).

Explain how algorithmic bias could arise in a given system

2 marks | Grade 6

1. Say what is wrong with the training data: who is missing or under-represented?
2. Link it to the result: the system works less well, or unfairly, for that group.
3. Use the people and the system named in the question.

Example. A voice assistant was trained mainly on recordings of adults from one region of the UK. Explain how this could lead to algorithmic bias. (2 marks)

Model answer. The system only learns from its training data, which contains few children or people with other accents (1), so it will misunderstand these users more often, treating them unfairly compared with other users (1).

Explain who is responsible when an AI or robot causes harm

3 to 4 marks | Grade 7

1. List the people or organisations who could be responsible.
2. Explain why the machine making its own decision makes blame unclear.
3. Mention machine learning or the 'black box' problem if it is relevant.
4. Link to the law: most laws were written for human decision-makers.

Example. A self-driving taxi with no human driver goes through a red light and injures a cyclist. Explain why it is difficult to decide who is legally liable. (3 marks)

Model answer. No human was driving, so the usual approach of holding the driver responsible does not work (1). Liability could lie with the car manufacturer, the software developers or the taxi company that runs and maintains the car (1). If the car uses machine learning, its behaviour came from its training data rather than a rule someone wrote, and it may not be possible to explain why it went through the light (1).

6-mark 'Discuss' the use of AI or robots

6 marks **Grade 8**

1. Start with the benefits in this context.
2. Cover the specification issues that apply: safety, algorithmic bias, accountability, legal liability.
3. Develop each point with a consequence in the context.
4. Conclude, e.g. whether a human should stay in control.

Example. A car manufacturer plans to sell fully self-driving cars that use machine learning. Discuss the ethical and legal issues of allowing these cars on public roads. (6 marks)

Model answer. Marked in levels: the top level needs a balanced answer with developed points applied to self-driving cars, and a conclusion. A strong answer might say: Self-driving cars do not get tired or distracted, so they could reduce crashes caused by human error, and they could give independence to people who cannot drive. However, safety is a concern: a car may meet a situation that was not in its training data, such as unusual roadworks, and make a dangerous decision. If the training data had few examples of, say, wheelchair users or children, the car could be worse at detecting them, which is algorithmic bias with serious consequences. After a crash it is unclear who is accountable and legally liable (the owner, the manufacturer or the software developers), and the car may not be able to explain its decision. Taxi and lorry drivers could also lose their jobs. Overall, the cars should only be allowed once they have been widely tested on representative data and the law makes clear who is liable in a crash.

Don’t lose marks

- Saying the AI 'chose to be biased': explain that it learned from unrepresentative or biased training data.
- Saying the robot or AI is to blame: a machine cannot be held legally responsible, so a person or company must be.
- Calling every computer program AI.
- Writing only about job losses when the question asks about ethical and legal issues: use the specification terms.
- Forgetting the benefits in a 'Discuss' answer.

More tips

Memory tricks

- The four specification issues: **SLAB**: Safety, Liability, Accountability, Bias.
- Bias in, bias out: an AI is only as fair as the data it learned from.
- Accountability = who must answer for it. Liability = who is legally responsible and may have to pay.
- Is it AI? If it learns or makes decisions a human would normally make, yes; if it just follows fixed steps like a calculator, no.

Exam technique

- Label each point with a specification term (safety, accountability, algorithmic bias, legal liability), then explain it in the context.
- For any bias question, mention the training data.
- A strong conclusion often recommends keeping a human in control of important decisions.

What each grade needs

What you need to be able to do, from the first marks up to the top grade.

Grade 3 — **Define artificial intelligence**
AI is a computer system that carries out tasks that normally need human intelligence, such as recognising speech.

Grade 3 — **Give uses of AI and robots**
For example, medical scan analysis, fraud detection, chatbots, self-driving cars and warehouse robots.

Grade 4 — **Describe machine learning**
A type of AI that learns patterns from training data instead of following rules written for every situation.

Grade 5 — **Explain how algorithmic bias happens**
If the training data is unrepresentative or reflects past prejudice, the system learns to treat some groups unfairly.

Grade 6 — **Explain safety risks of autonomous machines**
A self-driving car or robot that makes a wrong decision, meets a new situation or is hacked can injure people.

Grade 7 — **Explain accountability and legal liability problems**
It is hard to say who is to blame, or who must pay, when a machine made the decision itself.

Grade 8 — **Discuss using AI in a given context**
Weigh benefits such as speed and consistency against safety, bias, accountability and liability, then conclude.

Free notes, Full packs, practice papers and mock exams at chhetriacademy.com. Written for this site; not taken from past papers. Stuck? 1-to-1 tutoring with a free 30-minute trial lesson.