Teacher Answer Key

Case Study

1. Machine learning systems learn patterns from historical data. If the company's ten years of hiring data reflects a male-dominated workforce, the model will have learned that 'male candidate characteristics' correlate with being hired, and it will replicate that pattern in new recommendations. The bias came from the historical data, not from anyone explicitly programming a rule saying 'prefer men.'

2. The reading explains that in rule-based AI, someone wrote the discriminatory rule, making responsibility clear. In machine learning, the bias is embedded in historical training data, making responsibility more diffuse. In Jasmine's case, the company that chose to train on biased historical data and deployed the system without auditing for gender bias bears responsibility, even if no single programmer intentionally wrote a discriminatory rule.

Multiple Choice

1. B

2. B

3. C

Matching

1-D, 2-E, 3-A, 4-C, 5-B

Fill-in-the-Blank

1) algorithms, 2) subset, 3) training data

Short Answer (Sample Responses)

Q1: A traditional algorithm for spam filtering would require a programmer to write every rule explicitly — if the email contains 'click here to claim,' mark it spam. A machine learning model instead ingests millions of labeled emails and learns to detect spam patterns on its own, including patterns the programmer never thought to specify. The machine learning approach adapts as new spam tactics emerge simply by retraining on updated data.

Q2: The reading suggests asking whether the system is rule-based or machine learning, what data it was trained on, and whether that data was diverse, recent, and representative. Each matters because rule-based systems have human-authored biases that are traceable, while ML systems absorb biases from data that may

reflect historical inequalities — knowing which type you are dealing with determines how to evaluate its reliability and fairness.

Q3: In precise technical language, 'algorithm' refers specifically to a deterministic set of instructions, not necessarily a learning or adaptive system. Using it loosely to mean any automated decision system obscures whether the system follows fixed rules or learns from data, which matters enormously for understanding bias, accountability, and how the system can be audited or corrected.