Teacher Answer Key

Case Study

1. The facial recognition system's training data was not diverse enough; it contained predominantly lighter-skinned faces. The AI learned patterns from that data and became accurate at recognizing faces similar to what it had seen. Since Priya's face type was underrepresented in the training set, the model had not learned the patterns needed to reliably recognize her.

2. The developers should have ensured the training dataset included a diverse and balanced representation of skin tones, ages, genders, and ethnicities. They also should have tested the system's performance separately across demographic groups before deploying it in a school setting to identify accuracy gaps before they affected real students.

Multiple Choice

1. A

2. C

3. A

Matching

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

Fill-in-the-Blank

1) garbage, 2) unsupervised, 3) parameters

Short Answer (Sample Responses)

Q1: During training, the model makes a prediction on a piece of data, then the loss function measures how wrong that prediction was. Backpropagation then calculates how each parameter contributed to the error, and gradient descent adjusts those parameters slightly to reduce future errors. This cycle runs billions of times until the model's predictions become accurate.

Q2: Those early systems were trained on datasets that did not include enough examples of darker-skinned faces, so the AI never learned to recognize those facial features accurately. This reveals that training data must be diverse and representative; if the data does not include a group, the AI will perform poorly for that group, creating real-world harm.

Q3: Training a large model requires running the prediction-adjustment cycle billions of times across billions of data examples, which requires thousands of specialized

GPU chips working simultaneously. Even with that computing power, the sheer scale of the calculations needed to adjust trillions of parameters means the process takes weeks or months to complete.