1. Short Answer Questions

Q1: Define algorithmic bias and provide two examples of how it manifests in AI systems.
Answer:
Algorithmic bias occurs when an AI system produces unfair or discriminatory outcomes due to flawed assumptions, biased training data, or model design.

Examples:

1. **Gender bias in hiring tools:** Amazon's AI recruiting tool favored male candidates because it was trained on resumes from a male-dominated workforce.

2. **Racial bias in facial recognition:** Some facial recognition systems have higher error rates for darker-skinned individuals due to underrepresentation in training data.

Q2: Explain the difference between transparency and explainability in AI. Why are both important?
Answer:

- **Transparency** refers to openness about how an AI system is developed, including data sources, model architecture, and decision-making processes.

- **Explainability** is the ability to interpret and justify AI decisions in understandable terms for users.

Importance:

- **Transparency** builds trust and accountability.

- **Explainability** ensures users can challenge incorrect or biased decisions (e.g., loan denials).

Q3: How does GDPR (General Data Protection Regulation) impact AI development in the EU?
Answer:
GDPR imposes strict rules on AI development, including:

- **Right to explanation** (users can demand reasoning behind AI decisions).

- **Data minimization** (only necessary data can be collected).

- **Consent requirements** (explicit user permission for data processing).

- **Bias mitigation** (systems must avoid discriminatory outcomes).

2. Ethical Principles Matching

- **A) Justice** → Fair distribution of AI benefits and risks.

- **B) Non-maleficence** → Ensuring AI does not harm individuals or society.

- **C) Autonomy** → Respecting users' right to control their data and decisions.

- **D) Sustainability** → Designing AI to be environmentally friendly.