## Part 1: Theoretical Understanding (30%)

### Short Answer Questions

**Q1: Define algorithmic bias and provide two examples of how it manifests in AI systems.**

Algorithmic bias refers to systematic and unfair discrimination in the outcomes of AI systems, often caused by biased data or flawed model design.

Examples:

1. Facial recognition systems that perform poorly on darker-skinned individuals due to underrepresentation in the training data.
2. Hiring algorithms that favor male candidates over female ones because historical data reflects gender bias in previous hiring decisions.

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

Transparency refers to how openly the inner workings and decision-making process of an AI system are disclosed, including access to data sources, algorithms, and design choices.

Explainability focuses on how well a human can understand why an AI system made a specific decision or prediction.

Both are important because transparency builds trust by allowing oversight and auditing, while explainability helps users and stakeholders make sense of AI decisions, which is crucial for accountability and ethical use.

**Q3: How does GDPR (General Data Protection Regulation) impact AI development in the EU?**

GDPR impacts AI development by requiring explicit user consent for data collection and processing. It mandates data minimization and privacy by design, grants users the right to explanation when automated decisions affect them, and imposes strict penalties for non-compliance. This influences how AI systems are trained, deployed, and audited within the EU.

### Ethical Principles Matching

Ensuring AI does not harm individuals or society. – Non-maleficence
Respecting users' right to control their data and decisions. – Autonomy
Designing AI to be environmentally friendly. – Sustainability
Fair distribution of AI benefits and risks.