Part 1: Theoretical Understanding (30%)

1. Short Answer Questions

Q1: Algorithmic Bias - **Definition:** Algorithmic bias occurs when an AI system produces systematically unfair outcomes for certain groups due to biased data, design, or assumptions. - **Examples:** 1. Facial recognition software performing poorly on darker-skinned individuals due to underrepresentation in training data. 2. Hiring algorithms favoring male candidates because historical hiring data was biased toward men.

Q2: Transparency vs Explainability - **Transparency:** The extent to which the internal workings of an AI system are open and accessible. - **Explainability:** The ability to describe how and why an AI system reached a particular decision. - **Importance:** Transparency allows stakeholders to inspect the system for biases or errors, while explainability ensures that decisions can be understood and trusted by users.

Q3: GDPR Impact on AI Development - GDPR enforces data protection and privacy regulations in the EU. It impacts AI by: - Requiring explicit consent for data collection. - Granting users rights such as data access, correction, and deletion. - Encouraging transparency and explainability in AI decisions to comply with legal obligations.

---

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.