Part 1: Theoretical Understanding (30%)

1. Short Answer Questions

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

Algorithmic biases refer to systematic errors in machine learning, which result in unfair or discriminatory outcomes. According to Jonker & Rogers (2024), algorithmic bias reflects and reinforces unfair or discriminatory outcomes. Examples of algorithmic bias manifestations include a facial recognition system with high error rates for people of color, such as black women, or AI hiring tools that discriminate against specific population demographics, such as females.

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

Transparency and explainability in AI aim to make AI systems more understandable and trustworthy. AI-transparency involves openly sharing information about how an AI system is built and functions. Therefore, transparency promotes openness and honesty of AI systems, enabling people to know what occurs behind the scenes. On the other hand, AI explainability provides clear decisions for specific AI decisions. This allows people to understand AI decision-making processes. Overall, both help build customer trust with AI by improving transparency between businesses and customers regarding AI systems.

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

The GDPR addresses the handling of personal data, and the Artificial Intelligence Act, which concerns developing and deploying artificial intelligence systems (Mondoh, 2025). These regulations impact AI development in the EU as they require systems that process data to adhere to principles on a lawful basis for processing, data accuracy, and transparency. Also, it grants people

the right to access, rectify, object to, and delete their data. This can be challenging when implementing AI models. This regulation also limits automated decision-making with legal or significant effects, impacting AI applications in hiring or credit scoring.

2. Ethical Principles Matching

Match the following principles to their definitions:

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.*

References

Jonker, A., & Rogers, J. (2024). *What Is Algorithmic bias?* IBM; IBM. https://www.ibm.com/think/topics/algorithmic-bias

Mondoh, B. S. (2025). *Understanding the GDPR and EU AI Act: Key Insights for Businesses*. Thebarristergroup.co.uk; The Barrister Group. https://thebarristergroup.co.uk/blog/understanding-the-gdpr-and-eu-ai-act-key-insights-for-businesses

.