A4.4

Ethical considerations

CHAPTER ASSIGNMENT · 105 MARKS

Name _____ Class _____ No. _____ Date _____

Answer every question in the space given. Show your working. You may use the chapter page to help you.

Questions from *Computer Science for the IB Diploma*, 1st edition (2025), Paul Baumgarten, Ioana Ganea, Carl Turland (Hachette Learning), for classroom use.

Review Questions (p. 278)

1 An AI company develops a facial recognition system used in public surveillance.

(a) Outline three ethical implications of using facial recognition technology in public spaces.

(i) Identify two potential biases that could arise in facial recognition systems.

(ii) Outline the societal impacts of each.

(c) Outline two measures that could be implemented to address these ethical concerns and biases.

2 A social-media company uses algorithms to personalize newsfeeds based on user interactions.

(a) Outline two potential ethical issues related to algorithmic bias in personalizing newsfeeds.

(b) Outline two strategies the company could implement to ensure the ethical use of personalization algorithms.

(i) Identify two implications of lack of transparency in algorithmic decision-making.

(ii) Outline two methods to improve transparency in algorithmic decision-making.

3 A university uses AI to make admissions decisions based on application materials.

(a) Outline three potential ethical concerns with using AI in university admissions.

(i) Outline two possible biases that could arise in this AI system.

(ii) Outline their impact on students.

(c) Describe measures to address these ethical concerns and biases.

Linking Questions (p. 278)

1. How can machine learning be applied to optimize network traffic management? (A2)

2. How does database programming in SQL differ from programming computationally in a high-level language? (A3, B2)

3. To what extent are developments in machine learning ethical? (TOK)

4. How can larger models be processed using GPUs and cloud processing? (A1)

5. Can machine learning find and improve network security problems? (A2)

Exam Practice Questions — Topic A4 Machine Learning (pp. 279–280)

1 **Health monitoring app:** A tech startup has developed a health monitoring app that uses machine learning to predict potential health issues based on user-inputted symptoms, lifestyle data and historical health data. The app classifies user health into categories such as "low risk", "medium risk" and "high risk".

(i) State whether this system should be classified as artificial intelligence or machine learning. [1]

(ii) Outline one reason for your choice. [2]

(b) Describe the potential need for specialized hardware (e.g. GPUs) in deploying this app on mobile devices. [2]

(c) Describe the importance of data cleaning in this scenario, particularly addressing missing values in lifestyle data. [2]

(d) Describe how feature selection could impact the accuracy and efficiency of the predictive model used in the app. [2]

(e) Suggest the type of machine learning algorithm that would be suitable for this classification task. [4]

(f) Outline the implications of choosing a high value of k in a k-nearest neighbours (KNN) algorithm for this application. [2]

(g) Outline three ethical concerns related to privacy and data security in health-related apps. [3]

(h) Describe two measures that could be implemented to address potential biases in the data set, especially relating to underrepresented groups. [2]

2 **Autonomous public transport system:** A city plans to implement an autonomous bus service that uses machine learning to optimize routes based on traffic patterns, weather conditions and passenger demand.

(i) Define "edge computing". [1]

(ii) Describe its relevance in real-time data processing for autonomous vehicles. [2]

(b) Describe whether a deep learning model would be more effective than a traditional machine learning model for processing complex environmental data. [2]

(c) Describe how data normalization affects the performance of machine learning models dealing with varied data types such as weather conditions and traffic density. [2]

(i) Identify a common data quality issue that might arise with real-time traffic data. [1]

(ii) Outline a preprocessing step to mitigate these issues. [2]

(e) Describe how reinforcement learning could be applied to optimize bus routes dynamically. [2]

(f) Discuss the potential use of transfer learning from other cities' traffic management systems to improve route optimization. [2]

(g) Describe the ethical implications of using surveillance data (e.g. from traffic cameras) in training machine learning models for public transport systems. [2]

(h) Describe the societal impacts of replacing human-driven buses with autonomous buses, including job displacement and public safety. [2]

3 **AI-powered recruitment tool:** A multinational corporation implements an AI-powered tool to screen job applications and predict the suitability of candidates based on their résumés and answers to pre-interview questions.

(a) Describe the classification vs regression nature of the predictive model used by the AI tool. [2]

(b) Describe the impact of processing speed and memory requirements on the scalability of the AI tool across the corporation's global offices. [2]

(i) Identify potential biases in the training data set. [1]

(ii) Outline how these could be mitigated during data preprocessing. [2]

(d) Describe the importance of feature selection in improving the predictive accuracy of the AI tool. [2]

(e) Describe the use of a decision tree model over a regression-based model for this classification task. [3]

(f) Outline two ethical concerns related to AI decision-making in recruitment, particularly in terms of fairness and transparency. [2]

(g) Describe two methods to ensure the ethical use of AI in recruitment with respect to increasing transparency and accountability. [2]

4 **Retail customer segmentation:** A large retail chain uses machine learning to segment its customer base to personalize marketing strategies and improve customer service.

(a) Discuss whether supervised or unsupervised learning is more appropriate for customer segmentation. [4]

(b) Outline the potential benefits of using cloud computing resources over in-house servers for processing large customer data sets. [2]

(c) Outline the role of outlier detection in customer segmentation. [2]

(i) Identify a clustering algorithm suitable for handling large data sets with high dimensionality. [1]

(ii) Describe a reason for your choice. [2]

(e) Describe the potential privacy issues that may arise from the detailed segmentation of customers' buying habits. [3]

(f) Describe strategies to mitigate the risk of discriminatory marketing practices that could result from biased data in customer segmentation. [2]

5 **Natural-disaster prediction and management:** A government agency deploys machine learning models to predict natural disasters such as floods and earthquakes, aiming to enhance preparedness and response strategies.

(a) Describe the application of neural networks in predicting natural disasters and the kind of data they might process. [2]

(b) Describe the impact of using real-time data processing on system requirements and infrastructure. [2]

(c) Describe the challenges associated with integrating and cleaning data from multiple sources, such as satellite imagery and geological sensors. [2]

(d) Describe the role of data augmentation in improving the accuracy of predictions in areas with sparse historical data. [2]

(e) Describe the use of deep learning over traditional models for predicting complex natural disaster patterns. [2]

(f) Describe how machine learning models can be trained to adapt to new types of disaster data over time. [2]

(g) Describe the ethical implications of false positives and false negatives in disaster prediction models. [3]

(h) Describe protocols for data governance that may ensure sensitive geographical and personal data used in predictions are protected. [2]

6 **Automated cyberbullying detection system:** A software company is developing an automated system to detect and flag instances of cyberbullying on social-media platforms using natural language processing and machine learning.

(i) Define "natural language processing (NLP)". [1]

(ii) Describe its relevance in detecting cyberbullying. [2]

(i) Describe the computational challenges associated with processing large volumes of social-media data in real time. [2]

(ii) Outline appropriate hardware solutions. [2]

(c) Describe the potential preprocessing steps needed for textual data from social-media posts to prepare it for machine learning models. [2]

(d) Explain the importance of handling sarcasm and ambiguities in text when setting up preprocessing pipelines for detecting cyberbullying. [2]

(e) Describe the use of analytical rule-based systems vs machine learning models in the context of cyberbullying detection. [2]

(f) Outline the ethical considerations of implementing an automated cyberbullying detection system, particularly regarding false positives and false negatives. [3]

(g) Describe the potential privacy implications of analysing users' social-media content, even for the purpose of detecting cyberbullying. [2]

End of assignment · Check your answers against the chapter at gurujazzy.com/subjects/ib-computer-science/a4-4-ethical-considerations/