

「データサイエンス基礎」中間試験問題

数理・データサイエンス教育センター

2022年5月15日

Q1. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「(a) とは、データを分析することによって、分析対象に関する (b) を得ること。例えば、医療におけるがんの再発に関する知識を獲得すること。」

1. a. 知識発見、b. 未知の知見
2. a. 判断支援、b. 未知の知見
3. a. 活動代替、b. 既知の知見
4. a. 原因究明、b. 既知の知見

Q2. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. データサイエンスが発展した理由の一つとして、インターネットの発達によりビッグデータが入手可能になったことが挙げられる
- b. データサイエンスが発展した理由の一つとして、コンピュータの性能向上で大規模なデータが扱えるようになったことが挙げられる

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q3. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. メタデータとは、ある目的のために利用者自身で収集したデータのこと
- b. 2次データとは、他の目的のために事前に収集されていたデータのこと

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q4. 情報管理の3原則として、正しいものを選べ。

1. 完全性・機密性・可用性
2. 安全性・整合性・可用性
3. 安全性・秘密性・整合性
4. 完全性・秘密性・整合性

Q5. 次の記述中の空欄に入る適切な語句を選べ。

「きゅうり農家の直面した課題は、きゅうりの仕分けを、仕分けする人によらずに統一的な基準で行うことであった。正しく仕分けされたきゅうりの画像を多量に用意し、AI に画像と仕分けの等級の対応関係を学習させてこの問題を解決した。」この事例は AI を使った（ ）である

1. 判断支援
2. 活動代替
3. 仮説検証
4. 原因究明

Q6. 次の記述中の a と b の空欄を受ける適切な語句の組み合わせを選べ。

「人間は、一般に「(a)」によって直感的に判断しがちだが、(b) と呼ばれる先入観による偏りが伴う場合が多い」

1. a. 速い思考、b. 認知バイアス
2. a. 遅い思考、b. 認知バイアス
3. a. 速い思考、b. 確認バイアス
4. a. 遅い思考、b. 確認バイアス

Q7. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. 仮説検証とは、仮説が正しいかどうかをデータによって検証すること。
- b. 原因究明とは、ある出来事の原因（その出来事の発生に影響する要因）が何であるかを解明すること。

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q8. 次の記述中の a の空欄に入る適切な語句の組み合わせを選べ。

「ELSI とは、科学技術の発展が引き起こす (a) な問題」

1. a. 倫理的、法的、社会的
2. a. 倫理的、経済的、社会的
3. a. 心理的、法的、社会的
4. a. 倫理的、法的、思想的

Q9. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「多次元データの可視化は、データの間の (a) をなるべく変えないようにデータを 2 次元か 3 次元に低次元化してから可視化する。(b) などの手法がある」

1. a. 関係、b. 主成分分析
2. a. 配置、b. 主成分分析
3. a. 関係、b. 回帰分析
4. a. 配置、b. 回帰分析

Q10. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. データの可視化により、数量の比較、関係性、変化（推移）が分かりやすくなる。
- b. データの可視化はグラフを作ることが目的である

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q11. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. 画像認識では、黒人男性をゴリラとして分類する事件があったが、これはアルゴリズムバイアスの例である
- b. 人種・年齢・ジェンダーのような特徴はアルゴリズムバイアスが起りにくい

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q12. 次の記述中の空欄に入る適切な語句を選べ。

「くら寿司は目利き職人の代わりに AI システムを開発した。この AI はマグロの尾の切断面の画像をもとにマグロを最上級、上級、並の 3 等級に正しく分類するものである。」この事例は AI を使った（ ）である。

1. 判断支援
2. 活動代替
3. 仮説検証
4. 原因究明

Q13. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「匿名化とは、データの中から個人を特定できるデータ（個人 ID）を (a) し、特定の個人の情報であることが分からないように処理すること。連結可能匿名化とは、個人 ID を仮 ID に置き換え、

個人 ID は (b) で管理する方法。個人特定が必要な場合に限り、厳重に保管した (b) を使う。」

1. a. 削除、b. 連結表
2. a. 保留、b. 連結表
3. a. 削除、b. 参照表
4. a. 保留、b. 参照表

Q14. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. 他人のデータを勝手に使っても問題ない
- b. 自分のデータや文章を使い回しても問題ない

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q15. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. データを匿名化すれば個人情報の漏洩は絶対にない。
- b. データにパスワードをかければ、情報漏洩の心配は全くない。

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q16. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. 天気予報は予測である。
- b. ナビで示したルートはある基準で最適化されたものである。

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q17. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. ルールベース AI はあらかじめ人間が記述したルール通りに判断を行うので、暗黙知の判断に向いている。
- b. 機械学習型 AI は人間がルールを記述する必要がなく、データさえあれば開発できる。そのため、暗黙知の判断に向いている。

1. a. 正、b. 正

2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q18. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. 自分が書いた論文を英訳して再投稿しても不正ではない
- b. 教え子の卒論を微修正して第一著者として論文投稿しても不正ではない

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q19. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「PPDAC サイクルの A は (a) の A で、データを (a) して可視化したり、問題点を (a) します。明確な問題 (P) と計画 (P) の立案があつてこそ、実際のデータ収集 (D) や (a)(A) を行うことができ、(b) へ導くことができます。」

1. a. 分析 (Analysis)、b. 仮説の設定
2. a. 解答 (Answer)、b. 仮説の設定
3. a. 分析 (Analysis)、b. 問題解決
4. a. 解答 (Answer)、b. 問題解決

Q20. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「統計的差別とは、(a) な分析の結果として差別や不平等が (b) されてしまうこと」

1. a. 統計的、b. 肯定
2. a. 統計的、b. 否定
3. a. 科学的、b. 肯定
4. a. 科学的、b. 否定

Q21. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「AI の開発では、(a) が重要な意味を持つ。例えば、画像中の物体を認識する AI を開発するには、画像データに何が写っているかを (b) として付加する必要がある。数十万枚の画像に一つずつ手作業で (b) を付加するなど、アノテーションの作業には大変な労力が必要になる。」

1. a. アノテーション、b. ログデータ
2. a. データのクリーニング、b. ログデータ
3. a. アノテーション、b. メタデータ
4. a. データのクリーニング、b. メタデータ

Q22. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「PPDAC サイクルの D は (a) の D で、(a) の (b) を行います。必要なデータが揃っているかどうかのチェックが大事です。」

1. a. データ (Data)、b. 収集・整備
2. a. デジタル化 (Digital)、b. 収集・整備
3. a. データ (Data)、b. 推進
4. a. デジタル化 (Digital)、b. 推進

Q23. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

PPDAC サイクルの C は (a) の C で、分析した結果から (a) を導き出します。その後、分析レポートを作成し、分析した結果から (b) に繋がるアクションを提案します。

1. a. 結論 (Conclusion)、b. 価値ある施策
2. a. 結論 (Conclusion)、b. 仮説の設定
3. a. 分類 (Classification)、b. 価値ある施策
4. a. 分類 (Classification)、b. 仮説の設定

Q24. 次の a と b の記述の正誤として正しい組み合わせを選べ。

- a. デジタルトランスフォーメーション (Digital Transformation、DX) とは、デジタル技術でビジネスを変え、価値の創出や効率化を行うこと。
- b. 手続きのデジタル化もデジタルトランスフォーメーションの例である。

1. a. 正、b. 正
2. a. 正、b. 誤
3. a. 誤、b. 正
4. a. 誤、b. 誤

Q25. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「(a) とは、心理学者・行動経済学者の (b) が提唱した概念である。」

1. a. 「効率思考/経済思考」、b. P. Schelling
2. a. 「速い思考/遅い思考」、b. P. Schelling
3. a. 「効率思考/経済思考」、b. D. Kahneman
4. a. 「速い思考/遅い思考」、b. D. Kahneman

Q26. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「クリエイティブ・コモンズ・ライセンスとは、(a) の (b) を許可するライセンスのこと」

1. a. 著作物、b. 再利用
2. a. 著作物、b. 再発行

- 3. a. 制作物、b. 再利用
- 4. a. 制作物、b. 再発行

Q27. 次の記述中の a と b の空欄に入る適切な語句の組み合わせを選べ。

「PPDAC サイクルの2つ目の P は (a) の P で、研究 (a) の作成、既存のデータを使うのか、新たに調査を行ってデータを取得するのか、不足している (b) を行います。」

- 1. a. 計画 (Plan)、b. 知識の習得
- 2. a. 問題 (Problem)、b. 知識の習得
- 3. a. 計画 (Plan)、b. 研究費の獲得
- 4. a. 問題 (Problem)、b. 研究費の獲得

Q28. 次の記述中の a と b の空欄を受ける適切な語句の組み合わせを選べ。

「データ駆動社会とは、データに基づく「(a)」によって (b) を行う社会のこと」

- 1. a. 速い思考、b. 意思決定
- 2. a. 遅い思考、b. 意思決定
- 3. a. 速い思考、b. 条件反射
- 4. a. 遅い思考、b. 条件反射

Q29. 次の記述中の a と b の空欄を埋める適切な語句の組み合わせを選べ。

「データサイエンスとは、(a) から (b) を創造するアプローチのことである。」

- 1. a. データ、b. 万物
- 2. a. 資源、b. 万物
- 3. a. データ、b. 新たな価値
- 4. a. 資源、b. 新たな価値

Q30. 次の記述中の空欄に入る適切な語句を選べ。

「銀行業務の従来の課題として、投資商品のチェックに手間がかかることがあった。りそな銀行は投資商品に問題がないかどうかを自動的に判断する AI を導入した。」この事例は AI を使った（ ）である。

- 1. 判断支援
- 2. 活動代替
- 3. 仮説検証
- 4. 原因究明