

「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 の空欄に入る適切な語句の組み合わせを選べ。

「(a) とは、データ間の未知の関係性を発見すること。」

1. a. データ同化
2. a. クラスタリング
3. a. パターン発見
4. a. 汎化

Q2. 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。

「現在技術で実現可能な AI は (a) であり、(b) は実現不可能である」

1. a. 特化型 AI、b. 汎用型 AI
2. a. 強化型 AI、b. 汎用型 AI
3. a. 特化型 AI、b. 活用型 AI
4. a. 強化型 AI、b. 活用型 AI

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. 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

Q4. 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。

「2016 年、(a) データ活用推進基本法により、オープンデータが日本で初めて法律上で定義された。この法律では、国や自治体が保有する (a) データについて、(b) などを通じて容易に利用できる措置を講ずる、としている。」

1. a. 官民、b. インターネット
2. a. 官民、b. 官報などの公的メディア
3. a. 民間、b. インターネット
4. a. 民間、b. 官報などの公的メディア

Q5. 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

Q6. 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

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. 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

Q8. 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 サイクルの最初の P は (a) の P で、(a) の把握と明確化、そして分析すべきデータの決定と (b) を行います。」

1. a. 問題 (Problem)、b. 仮説の設定
2. a. 計画 (Plan)、b. 仮説の設定
3. a. 問題 (Problem)、b. データの可視化
4. a. 計画 (Plan)、b. データの可視化

Q9. 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. 捏造とは、存在しないデータを作り出すこと
- b. 剽窃とは、データを都合よく書き換えること

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

Q10. 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。

「データの「可視化」はただの「作図」ではない。数量の比較、関係性、変化（推移）が分かりやすくなり、データの (a) に気づき、データ分析の (b) を見出すことができる。」

- 1. a. 特徴、b. 端緒
- 2. a. 特徴、b. 弱点
- 3. a. 長所、b. 端緒
- 4. a. 長所、b. 弱点

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. 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

Q12. 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. 個人情報とプライバシーははっきりした区別がある
- b. プライバシーは法的に保護されていない

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

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。

「SDGs は (a) の目標、169 のターゲットから構成され、『(b) の誰一人取り残さない』ことを誓っています。」

- 1. a. 17、b. 地球上
- 2. a. 27、b. 地球上

3. a. 17、b. 火星上

4. a. 27、b. 火星上

Q14. 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。

「データカタログサイト DATA.GO.JP によると、日本政府は、公共データを広く公開することにより、(a) の向上、企業活動の活性化等を通じ、我が国の社会経済の発展に寄与する観点から、(b) に適したデータ形式を、営利目的も含めた二次利用が可能な利用ルールで公開する「オープンデータ」の取組を推進している。」

1. a. 国民意識、b. インターネットでの流通

2. a. 国民意識、b. 機械判読

3. a. 国民生活、b. インターネットでの流通

4. a. 国民生活、b. 機械判読

Q15. 次の記述中の a、b の空欄に入る適切な語句の組み合わせを選べ。

「忘れられる権利とは、インターネット上のデータについて、(a) 保護のために削除やアクセスの遮断を (b) 権利」

1. a. プライバシー、b. 求める

2. a. 人身安全、b. 求める

3. a. プライバシー、b. 断る

4. a. 人身安全、b. 断る

Q16. 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

Q17. 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。

「オープンデータとは、(a) 上で扱いやすいデータ形式で、2 次利用が (b) なデータのこと」

1. a. コンピュータ、b. 可能

2. a. コンピュータ、b. 不可

3. a. インターネット、b. 可能

4. a. インターネット、b. 不可

Q18. 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

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 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

Q20. 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. SNS などで行うネット調査において対象を制限しない場合、データバイアスは生じない
- b. 授業アンケートで、自発的に提出した学生のみを対象とした場合、データバイアスが生じる

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

Q21. 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

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。

「ニューラルネットワークは (a) の手法の一つであり、(b) のネットワークの上で演算を行う」

1. a. 機械学習、b. 層状
2. a. エキスパートシステム、b. 層状
3. a. 機械学習、b. 線状

4. a. エキスパートシステム、b. 線状

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。

「データサイエンススキルは、(a)、データへの理解・検証、(b) などが含まれています。」

1. a. 数理・統計学、b. 機械学習の技法
2. a. 物理・生物学、b. 機械学習の技法
3. a. 数理・統計学、b. ロボットの製作
4. a. 物理・生物学、b. ロボットの製作

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. 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

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。

「データ解析とは、データに対して (a) 的な手法を適用することにより、データを可視化するだけでは分からない新たな知見を得ることです。データ解析の流れは、1. データの入手、2. データの確認、3. (c)、4. データの可視化、5. 記述統計による分析、6. 推測統計による分析、7. 結論の 7 つのステップから構成されます。」

1. a. 統計学、b. データの前処理
2. a. 解析学、b. データの前処理
3. a. 統計学、b. データの追加
4. a. 解析学、b. データの追加

Q26. 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

Q27. 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. Youtube 上のデータは構造化データである。
- b. ヤフーニュースに掲載された内容は全てテキストデータである。

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

Q28. 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・ライセンスが普及する前は、(a) や (b) により、データの再利用が難しかった」

- 1. a. 著作権法、b. 知的所有権法
- 2. a. 憲法、b. 知的所有権法
- 3. a. 著作権法、b. 民法
- 4. a. 憲法、b. 民法

Q29. 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. 人工知能を意味する AI とは、Artificial Ingredient の略である
- b. 人工知能には明確な定義がある

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

Q30. 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。

「『エビデンス ((a)) に基づく政策立案』において、エビデンスの抽出に手間がかかる問題を、行政のビッグデータから (b) を抽出することで解決したのが、神奈川県葉山町の事例である。」

- 1. a. 証拠や根拠、b. 無駄な予算
- 2. a. 証拠や根拠、b. 仮説
- 3. a. 事例と説明、b. 無駄な予算
- 4. a. 事例と説明、b. 仮説