



[特別付録]

# 模擬試験192問 問題

### 問1

ニューラルネットワークにおいて、勾配消失問題が起こりやすくなる条件として、最も適切な選択肢を1つ選べ。

1. 学習データの量が多い。
2. ネットワークのニューロン層の数が多い。
3. 学習データのサイズが揃っていない。
4. 活性化関数の微分値が1より大きい。

### 問2

2018年にGoogle社によって開発された自然言語処理モデルBERTについて述べた文章として、最も適切な選択肢を1つ選べ。

1. ALBERTをもとに改良されたモデルである。
2. 文章を「過去から現在」と「未来から現在」の双方向の情報から学習する。
3. エンコードとデコードにRNNが使われている。
4. 偽物のブログの作成等に使われる懸念があり、APIを通した利用申請が必要である。

### 問3

単語や記号同士の意味関係を認識できたとしても、それらが表す実世界の概念と結び付けられないという人工知能分野の問題の名称として、最も適切な選択肢を1つ選べ。

1. おもちゃの問題
2. トロツコ問題
3. シンボル・グラウンディング問題
4. フレーム問題

### 問4

入力層と出力層の2つの層から構成されるニューラルネットワークの名称として、最も適切な選択肢を1つ選べ。

1. ネオコグニトロン
2. 完全畳み込みネットワーク
3. 単純パーセプトロン
4. バイナリネットワーク

### 問5

畳み込みニューラルネットワーク(CNN)において、全結合層の役割として、最も適切な選択肢を1つ選べ。

1. 画像データ上の特徴を抽出して特徴マップを作る。
2. 画像のカテゴリ分類を行う。
3. 特徴マップの周囲を0で埋めることで、画像の端の特徴も考慮されるようにする。
4. 物体のわずかな位置変化による影響を下げ、過学習を抑制する。

### 問6

機械学習モデルのハイパーパラメータとして、最も不適切な選択肢を1つ選べ。

1. ランダムフォレストを構成する弱学習器の数
2. ニューラルネットワークの中間層の重み
3. ニューラルネットワークの学習率
4. K-Means においてクラスターの数意味する K の値

### 問7

ドロップアウトの性質について、最も適切な選択肢を1つ選べ。

1. ミニバッチごとにデータの正規化を行うことで過学習を防ぐ。
2. ランダムでユニット間の結合を削除することで過学習を防ぐ。
3. ランダムに一部のニューロンを無効化することで、汎化性能を上げる。
4. モデルの精度が向上しなくなった時点で訓練を打ち切ることで過学習を防ぐ。

### 問8

X社が「機能Aを有するプログラム」の特許を日本でのみ取得している。この状況について言えることとして、最も適切な選択肢を1つ選べ。

1. 日本国内のY社がこのX社の技術レベルを把握する目的で、特許記載のプログラムを自ら実装し検証実験を行った。Y社のこの行為は、X社の特許権を侵害する。
2. 日本国内のY社がX社からこのプログラムを購入し、Y社自らの業務に活用した。Y社のこの行為は、X社の特許権を侵害する。
3. 特定の企業に属さない個人W氏がこのX社の特許記載のプログラムを自ら実装し、個人の学習という私的な目的で使用した。個人としてのW氏のこの行為は、X社の特許権を侵害する。
4. 日本国外のZ社が、このX社の特許記載のプログラムを日本国外で実装し、日本国外で販売した。Z社のこの行為は、X社の特許権を侵害しない。

### 問9

リカレントニューラルネットワーク(RNN)を用いた自然言語処理において、入力系列長と出力系列長が一致しなくても対応できるように空白文字(blank)を加えて処理する技術として、最も適切な選択肢を1つ選べ。

1. CTC (Connectionist Temporal Classification)
2. CEC (Constant Error Carousel)
3. LSTM (Long Short Term Memory)
4. GRU (Gated Recurrent Unit)

### 問10

最初に開発された GAN (敵対的生成ネットワーク)に畳み込み層を取り入れることによって、学習を安定させ、より自然な画像を生成できるようになった GAN の改良モデルとして、最も適切な選択肢を1つ選べ。

1. chatGAN
2. DistilBERT
3. DCGAN
4. GNMT

### 問11

Transformer モデルに関する記述として、最も不適切な選択肢を1つ選べ。

1. BERT や GPT の大規模言語モデルは Transformer から構成されている。
2. Transformer はデータ処理の並列化が可能である。
3. Transformer は、長い系列の解析を得意とする LSTM を Source-Target Attention で橋渡しした構造である。
4. Transformer を取り入れたニューラル機械翻訳は長文に対する翻訳精度が従来モデルに比べて向上した。

### 問12

音声認識において、時間的に連続なアナログデータを、コンピュータで処理可能なデジタルデータに変換する A-D 変換について説明した文章として、最も適切な選択肢を1つ選べ。

1. 一定の時間間隔で信号を抽出することを量子化という。
2. 一定の時間間隔で信号を抽出することを標本化という。
3. 連続値を離散的な数値データに変換することを符号化という。
4. 0 と 1 で表されるデジタルデータに変換することを量子化という。

### 問13

以下の文章を読み、空欄に入る用語として最も適切な選択肢を1つ選べ。

AIの原理をいったん理解してしまうと、「これは知能ではなく、単なる自動化に過ぎない」とってしまう人間の心理を一般的に( )と呼ぶ。

1. AI 汎化
2. AI 支配
3. AI 効果
4. AI 現象

### 問14

個人情報保護法における個人識別符号について、最も適切な選択肢を1つ選べ。

1. DNA のデータなど、特別な技術を用いて加工しなければ個人を特定できないような情報は個人識別符号に含まれない。
2. 基礎年金番号は、個人識別符号に含まれる。
3. 個人識別符号は、数値化されている情報であることが必要であるため、指紋は、個人識別符号に該当しない。
4. 他の個人情報と照らし合わせることで特定の個人を識別できる情報を個人識別符号と定義される。

### 問15

音声認識のために「検定に合格した」を音素に分割したものとして、最も適切な選択肢を1つ選べ。

1. 検定に / 合格した
2. 検定 / に / 合格し / た
3. け / ん / て / い / に / ご / う / か / く / し / た
4. k/e/n/t/e/i/n/i/g/o/u/k/a/k/u/sh/i/t/a

### 問16

オントロジーの概念間の関係の1つである「part-of の関係」について、(ア) part-of (イ)の組み合わせとして、最も適切な選択肢を1つ選べ。

1. (ア)日本      (イ)関東
2. (ア)車輪      (イ)トラック
3. (ア)PC      (イ)スマートフォン
4. (ア)猫      (イ)動物

### 問17

以下の文章を読み、(ア)(イ)(ウ)の組み合わせとして、最も適切な選択肢を1つ選べ。  
敵対的生成ネットワークは、2種のニューラルネットワーク(ア)と(イ)から構成される。(ア)に(ウ)を入力することで、目標とするデータに近づけるようにマッピングを行う。(ア)が生成した画像を(イ)がもとの画像と識別しようとする。(ア)と(イ)を競合させるように訓練を行うことで、各々の性能を強化し、本物と見分けられないような新しい画像サンプルを生成する仕組みである。

- |                 |              |            |
|-----------------|--------------|------------|
| 1. (ア)ジェネレータ    | (イ)ディスクリミネータ | (ウ)ランダムノイズ |
| 2. (ア)エンコーダ     | (イ)デコーダ      | (ウ)本物のデータ  |
| 3. (ア)ディスクリミネータ | (イ)デコーダ      | (ウ)ランダムノイズ |
| 4. (ア)エンコーダ     | (イ)ディスクリミネータ | (ウ)本物のデータ  |

### 問18

無償でダウンロードし、自由に活用できるデータとして、最も適切な選択肢を1つ選べ。

1. 社員名簿データ
2. POS データ
3. オープンデータ
4. 定型データ

### 問19

以下の文章を読み、( )に入る選択肢として最も適切な選択肢を1つ選べ。  
強化学習の手法の1つはQ学習(Q-Learning)である。Q学習では、各時刻における状態とエージェントの行動とそれに対する評価にもとづいて( )の値を更新しつづける。

1. スラック関数
2. 行動価値関数
3. カーネル関数
4. 勾配関数

## 問20

AI 技術の利活用はユーザーに多くの利益をもたらす一方で、個人情報の流出や濫用といったリスクの原因にもなる。新サービスの導入・設計にあたって、プライバシー侵害の可能性を事前に予測し、仕様設計の段階から防止策を組み込もうとする考え方として、もっとも適切な選択肢を 1 つ選べ。

1. プライバシー・バイ・デザイン
2. プライバシー・バイ・プロダクト
3. プライバシー・バイ・エンジニアリング
4. プライバシー・バイ・メカニクス

## 問21

以下のうち、著作権法によって著作物として保護される可能性のあるものとして、最も不適切な選択肢を 1 つ選べ。

1. システムのユーザー向けに作成したマニュアルや説明書。
2. システムを稼働するためのコンピュータプログラム。
3. AI アルゴリズムを実装したのに使用したプログラミング言語。
4. 最新の AI 技術が発表されたという事実を記述し、さらにその技術について有識者が解説したニュース記事。

## 問22

RNN は時系列データの予測に使用される。時系列データは時間軸に沿って入力され、時間軸に沿って誤差が反映されると同時に RNN では、通常ニューラルネットワークと同じく誤差逆伝播法に従う。このように、通常の誤差逆伝播に加えて、時間軸に沿って誤差を反映する RNN の仕組みとして、最も適切な選択肢を 1 つ選べ。

1. BPTT
2. ELMo
3. 入力重み衝突
4. 出力重み衝突

### 問23

自然言語処理によって解決できるタスクについて、最も不適切な選択肢を 1 つ選べ。

1. 固有表現抽出タスクは、文章から人名や組織名などの固有表現を抽出できるため、プライバシー保護の観点からその使用を制限する動きがある。
2. 文章要約タスクは、記事の内容からタイトルを自動生成するのに使われる。
3. 文章分類タスクは、ウェブニュース記事のカテゴリ化に使われる。
4. 感情分析タスクは、1 度に複数の感情の種類についてスコアを出力することができる。

### 問24

過学習に関する説明として、最も不適切な選択肢を 1 つ選べ。

1. 特徴量の数が少ないのに、データの数が多いときに過学習になりやすい。
2. 過学習しやすいのは、バイアスが低くてバリエーションが高いときである。
3. 正則化は過学習を抑えるための手法の 1 つである。
4. 訓練データに対するモデルの精度よりもテストデータに対するモデルの精度の方が低いときに過学習になっている可能性がある。

### 問25

(ア)(イ)に最もよくあてはまる組み合わせを 1 つ選べ。

ここ 20 年ほど続いてきた第 3 次 AI ブームは(ア)で代表されると言われている。その中でインターネットの発展とともに分析対象となるデータ量が増えることで特に活用が進んだ技術として(イ)が挙げられる。

- |                 |            |
|-----------------|------------|
| 1. (ア)機械学習      | (イ)データの可視化 |
| 2. (ア)ディープラーニング | (イ)自然言語処理  |
| 3. (ア)推論・探索     | (イ)チェス AI  |
| 4. (ア)GPU       | (イ)強い AI   |

### 問26

2000 年以降に用いられる LSTM (Long Short-Term Memory) において、伝播する情報を制御するためのゲートとして、最も不適切な選択肢を 1 つ選べ。

1. 忘却ゲート
2. 入力ゲート
3. 出力ゲート
4. リセットゲート



### 問27

以下の文章を読み、(ア)(イ)に入る組み合わせとして最も適切な選択肢を1つ選べ。  
深層強化学習の手法である DQN では、(ア)の関数近似に(イ)を使用する。

1. (ア)行動価値関数      (イ)ボルツマンマシン
2. (ア)損失関数      (イ)Actor-Critic
3. (ア)行動価値関数      (イ)CNN
4. (ア)損失関数      (イ)BiRNN

### 問28

以下の文章を読み、(ア)(イ)に入る組み合わせとして、最も適切な選択肢を1つ選べ。  
勾配降下法を用いてニューラルネットワークの最適化を行う際に、(ア)である学習率を適切に設定しなかった場合、(イ)に陥ってしやすくなる。

1. (ア)重み      (イ)極大点
2. (ア)ハイパーパラメータ      (イ)大域最適解
3. (ア)重み      (イ)プラトー
4. (ア)ハイパーパラメータ      (イ)局所最適解

### 問29

EC サイトなどでは顧客の購買を促すために、その顧客が好みそうな商品を判断し推薦する手法が行われている。商品の特徴から推薦を行う手法として、最も適切な選択肢を1つ選べ。

1. 協調フィルタリング
2. 商品ベースフィルタリング
3. 行動ベースフィルタリング
4. 内容ベースフィルタリング

### 問30

不正競争防止法上、限定提供データに関する記述として、最も適切な選択肢を1つ選べ。

1. 特許法には保護されないが、創作性が認められる場合、著作権法による保護が可能である。
2. 限定提供データとして保護されるためには秘密として管理されている必要がある。
3. 限定提供データの一例は営業秘密である。
4. サービスや製品の改善を目的に企業間で共有されるデータが限定提供データに該当する。

### 問31

ロボットの制御に用いられる機械学習の手法について、最も適切な選択肢を 1 つ選べ。

1. 学習の成果を明確な基準にしたがって評価するために、教師データが付与されている教師あり学習が望ましい。
2. 行動の 1 ステップずつをリードする役割を果たす教師データを作るコストが大きいため、半教師あり学習が望ましい。
3. 特定の環境におけるロボットの複雑な行動を動的で表現するために深層強化学習が望ましい。
4. ロボットの挙動の不正常やハードウェアの故障を早めに検知したいため、異常検知によく用いられる教師なし学習が望ましい。

### 問32

クラウド上のストレージサービスの望ましい用途として、最も適切な選択肢を 1 つ選べ。

1. 容量の少ないデータを一時的に保存するため。
2. 個人情報を含む機密性の高いデータを保存するため。
3. クラウド上のサービスから出力される大量なログデータを保存するため。
4. インターネットに接続しない端末で利用したいデータ。

### 問33

空欄( )に入る内容として、最も適切な選択肢を 1 つ選べ。

ニューラルネットワークの学習において、( )ことで過学習を防ぐ。

1. ランダムで選び出した一部の隠れ層を無効化する。
2. 出力層にソフトマックス関数を位置づける。
3. ランダムに選び出した一部の重みをゼロにする。
4. 隠れ層の活性化関数を ReLU 関数にする。

### 問34

日本の特許法に基づく、プログラムの特許の出願に関して、最も適切な選択肢を1つ選べ。

1. X社は開発したプログラムを論文発表した。この時点で技術は公知となるが、「新規性の喪失の例外」の規定に基づいて、本プログラムの特許出願をすることは可能である。
2. X社は、開発したプログラムの特許として出願した。特許が登録される前に、本プログラムをオープンソースとして公開すると、技術は公知となり、新規性が失われる。
3. X社は、開発したプログラムの特許として出願した。特許法は日本国内にのみ適用されるので、出願前に日本国外の学会で同じ技術が発表されていたことが判明しても、新規性は喪失しない。
4. X社は、開発したプログラムの特許として出願した。Y社も直後に同じ技術を開発し、X社の特許が公開される前に特許出願した。Y社の出願時点でX社の特許技術は公知ではなく、Y社特許は認められる。

### 問35

(ア)に最もよくあてはまる選択肢を1つ選べ。

関数  $f(x) = 2x^3 - 2x + 6$  の  $x = 2$  における傾きは(ア)である。

1. 20
2. 18
3. 28
4. 22

### 問36

データあるいはモデルを知的財産として保護する場合の留意点として、最も不適切な選択肢を1つ選べ。

1. 機械学習モデルの学習用に入手したデータを加工し、体系的に構成したデータセットは原則として著作権法上の著作物とは見なされない。
2. 特定の条件を満たすものに提供先を限定しているデータは不正競争防止法上、限定提供データとして保護の対象になりうる。
3. 学習済みモデルを営業秘密として保護するためには、暗号化や難読化などの処理を施すことが望ましい。
4. モデルの学習のために収集した生データは著作物として保護されない。

### 問37

機械学習のためのサービスである Kaggle について、最も適切な選択肢を 1 つ選べ。

1. 強化学習を主テーマとする人工知能の学会
2. 機械学習のコンペティションである。
3. 研究論文をアップロードし公開するサイトである。
4. 大きな画像データのオープンデータベース

### 問38

著作権の利用あるいは著作権者の権利について、最も不適切な選択肢を 1 つ選べ。

1. 著作物の私的目的でのダウンロードは私的複製にあたるため、著作権者からの許諾を必要としない。
2. 著作権法に基づいて、著作権者による権利の 1 つは許諾であり、これは他の人に当該著作物に対する著作権を譲ることである。
3. 著作物が無断で第三者に利用されたときに著作権者は差し止めることができる。
4. 著作物を用いて学習用データを作成する際に、一定の要件のもとで著作者から許諾を得ずに行うことができる。

### 問39

自己回帰モデル(AR)を用いて行うことができる分析の対例として、最も不適切な選択肢を 1 つ選べ。

1. 来週の降雨量を予測する。
2. 株価の変動パターンを予測する。
3. 来年の失業率を予測する
4. スпамメールかどうかを予測する。

### 問40

空欄(ア)(イ)に入る組み合わせとして、最も適切な選択肢を 1 つ選べ。

過学習を抑制する工夫を取り入れながらモデルを学習することによって、(ア)が大きくなり、(イ)が小さくなる。

1. (ア)バリエーション (イ)バイアス
2. (ア)エラー (イ)ノイズ
3. (ア)ノイズ (イ)エラー
4. (ア)バイアス (イ)バリエーション

#### 問41

以下の文章を読み、(ア)に入る名称と(イ)に入る状況の組み合わせとして、最もよくあてはまる選択肢を1つ選べ。

「現代版バベルの塔」と呼ばれる人間のすべての知識を取り込むことを目指す野心な AI プロジェクト(ア)は、現在は(イ)。

- |                  |                         |
|------------------|-------------------------|
| 1. (ア)Cyc プロジェクト | (イ)自動応答の精度が最高水準に達している   |
| 2. (ア)東口ボくん      | (イ)全国高校生模試で高い偏差値を獲得している |
| 3. (ア)Cyc プロジェクト | (イ)未完成のまま継続されている        |
| 4. (ア)東口ボくん      | (イ)クイズ番組で優勝を果たしている      |

#### 問42

不正競争防止法における営業秘密に関する説明として、最も不適切な選択肢を1つ選べ。

1. 営業秘密を保護するために、知的財産権としての条件を満たしている必要がない。
2. 営業秘密として保護するためには、保護の対象物が暗号化などにより秘密管理されている必要がある。
3. 学習済みモデルを営業秘密として保護するためには、そのモデルが新規な発明として特許出願されていることが必須要件である。
4. 財務データ、顧客リスト、技術開発のノウハウはすべて営業秘密に該当する。

#### 問43

Python の特徴を表した記述として、最も不適切な選択肢を1つ選べ。

1. 数値計算、機械学習、データ処理などのためのライブラリは数多く Python 言語で提供されている。
2. C 言語や FORTRAN などよりも高速なプログラミング言語であるため、大量なデータの高速な処理が得意である。
3. Python はスクリプト言語というプログラミング言語のカテゴリに属する。
4. Python は Google Colaboratory の上で記述できる。

#### 問44

A3C の性質として、最も不適切な選択肢を 1 つ選べ。

1. 強化学習の課題の 1 つである報酬の揺らぎを軽減する効果がある。
2. 経験の自己相関を減らすことができる。
3. 学習の安定性と引き換えに、速度が比較的遅いアルゴリズムである。
4. 共有ネットワークと重みの値を同期しながら学習を行う。

#### 問45

以下の文章を読み、( ) に最もよくあてはまる選択肢を 1 つ選べ。

欠損値を適切な処理で埋めることも特徴量エンジニアリングの一部である。( ) 場合、回帰を利用して欠損値を埋める値を推測することができる。

1. 特徴量の種類が多い。
2. データ全体で欠損が発生する特徴量が 1 つのみである。
3. 欠損値の割合が全データに対しておよそ 2 割以下である。
4. 欠損が発生している特徴量と他の特徴量の間に強い相関が存在する。

#### 問46

以下の文章を読み、空欄(ア)(イ)に入る組み合わせとして、最も適切な選択肢を 1 つ選べ。

ニューラルネットワークの学習にドロップアウトを使うとする。ある特定のノードについて  $\alpha$  という割合でランダムに選択し無効にしたとする。この場合、モデルが予測結果を出力する際に、当該ノードに(ア)を(イ)する。

1. (ア)  $1 - \alpha$  (イ) 乗算
2. (ア)  $1 + \alpha$  (イ) 加算
3. (ア)  $\alpha$  (イ) 加算
4. (ア)  $2 \alpha$  (イ) 乗算

#### 問47

開発したプログラムを日本で特許出願するにあたり、発明者となれる人として最も適切な選択肢を1つ選べ。なお、特許法第33条に規定された「特許を受ける権利の移転」はなされていないものとする。

1. プログラムの具体的着想をしたが、プログラミング作業を行う者に指示を与えたのみで、プログラミング作業自体は行っていない者。
2. プログラムの具体的着想には関与していないが、プログラミング作業を行った者。
3. プログラムの具体的着想には関与していないが、そのプログラムの開発に資金や設備を提供した者。
4. プログラムの具体的着想には関与していないが、その着想者やプログラミングを行うエンジニアの業務管理を行う上司。

#### 問48

探索・推論の用途として最も適切な選択肢を1つ選べ。

1. 掃除ロボット
2. 洗剤の自動投入洗濯機
3. 自動運転車
4. 自動音声認識

#### 問49

以下の文章を読み、(ア)に最もよくあてはまる選択肢を1つ選べ。

教師なし学習において、未知の集合をいくつかのグループに分類するタスクは(ア)と呼ばれる。

1. バギング
2. クラスタリング
3. ブースティング
4. プーリング

### 問50

モラベックのパラドックスに関する記述で、最も適切な選択肢を 1 つ選べ。

1. AI がすべての可能性を考慮しようとするあまり、フリーズしてしまい、結局何もできなくなるという矛盾。
2. 特徴量と呼ばれる予測の手がかりを増やすことで逆に精度が下がることがある。
3. 「人工の知能」と名付けられているわりには、単独のアルゴリズムのそれぞれができるタスクは範囲が狭く、汎用的な役割を期待できない。
4. 幼児が本能に基づいて身につける知覚や運動能力を得ることは、AI による高度な計算に基づく予測よりも難しい。

### 問51

Inception モジュールから構成された CNN モデルとして、最も適切な選択肢を 1 つ選べ。

1. LeNet
2. U-Net
3. ResNet
4. GoogLeNet

### 問52

第 2 次 AI ブームの代表技術であるエキスパートシステムに関して、最も適切な選択肢を 1 つ選べ。

1. エキスパートシステムは、知識やルール間の整合性を担保できなかったため、実世界での利用に到らず研究に終わった。
2. 遺伝性の血液疾患を診断する Mycin（マイシン）は一般の内科医の診断精度を超えることができたものの、専門医の診断精度に到達できなかった。
3. ELIZA は心理カウンセラーの知識を持つエキスパートシステムである。
4. 専門分野の仕事に就く人間が失業する懸念により、エキスパートシステムの使用が制限された。



### 問53

人工知能(AI)をレベル1からレベル4までの4つのレベルに分類できる。複数レベルのAIが混在するシステムもある。このうち、レベル1のAIの説明として、最も適切な選択肢を1つ選べ。

1. 単純な自動制御プログラムであり、制御工学などの学術分野に対応する。
2. 特徴量の作成を人間に頼らず、自ら見出すことができる。
3. 探索や推論を活用しながら答えを出せる。
4. 大量なデータに基づいて予測する仕組みを持つ。

### 問54

オープンデータセットの利用に関して、最も不適切な選択肢を1つ選べ。

1. 機械判読が可能なファイル形式で提供される。
2. バイアスや異常値が著しく少ない。
3. 二次利用が可能。
4. 比較的大規模なデータを入手できることが多い。

### 問55

機械学習やディープラーニングの研究開発を支援する無償のプラットフォームのうち、ウェブ上で学術論文の検索と閲覧に利用できるものとして、最も適切な選択肢を1つ選べ。

1. NeurIPS
2. セメンティック Web
3. IEEE
4. Google Scholar

### 問56

データの次元の数が多くなるにつれて計算量が指数関数的に増えてしまうことを一般的に次元の呪いと呼ぶ。ここでいう次元の例として、最も適切な選択肢を1つ選べ。

1. 入力データの件数
2. 入力データにおける最大値
3. 入力データの特徴量の数
4. 入力データにおける単位

### 問57

特徴抽出を行う S 細胞層と位置ズレを許容するためにプーリング処理を行う C 細胞層の組み合わせを特徴とするモデルは、現在の畳み込みニューラルネットワーク (CNN) の原型であると思われる。このモデルの名称として、最も適切な選択肢を 1 つ選べ。

1. AlexNet
2. ネオコグニトロン
3. 多層パーセプトロン
4. オートエンコーダ

### 問58

不正競争防止法における営業秘密に該当するものとして、最も不適切な選択肢を 1 つ選べ。

1. 仕入れ先リスト
2. 商品の設計図
3. 個々の社員の生年月日
4. 実験データ

### 問59

LSTM (Long Short-Term Memory) において、出力重み衝突を防ぐための仕組みとして、最も適切な選択肢を 1 つ選べ。

1. CEC
2. 出力ゲート
3. 更新ゲート
4. フィードバック機構

### 問60

AI モデルを訓練する際に大量に電力が消費されることが近年問題視されている。電力消費を抑える工夫を取り入れた自然言語処理モデルとして、最も適切な選択肢を 1 つ選べ。

1. GPT-3
2. GPT-2
3. Switch Transformer
4. BERT

### 問61

2018年に施行された一般データ保護規則(GDPR)について、最も適切な選択肢を1つ選べ。

1. GDPRにおいて、欧州連合(EU)域外へ個人データの持ち出しに関しては、世界の全ての国に等しい規制を課している。
2. EUが策定した規定ではあるが、EU域外のデータに適用されることもある。
3. GDPRでは個人データの保護を強化するためにデータポータビリティを禁止している。
4. GDPRが適用されるためには、対象とする企業がEU域内に拠点があることが必要である。

### 問62

人工知能(AI)に関する説明として、最も正しい選択肢を1つ選べ。

1. 人工知能は、単純な条件分岐にしたがって動作することがある。
2. 人工知能とは自律的に動作する機械のことである。
3. 人工知能を構築することは、学習データを必須とする。
4. 人工知能を完全自動化した仕組みが機械学習である。

### 問63

以下の文章を読み、( )に入るものとして、最も適切な選択肢を1つ選べ。

リカレントニューラルネットワーク(RNN)において、時間軸に沿って伝播する重みの更新と一般的なニューラルネットワークとしての重みの更新の間に( )が起きやすい。

1. 勾配消失問題
2. 局所最適解問題
3. 重み重複問題
4. 重み衝突問題

#### 問64

以下の文章を読み、(ア)(イ)に入る組み合わせとして、最も適切な選択肢を1つ選べ。  
プログラムにおける記述の選択肢が広ければ広いほど、プログラムに(ア)が認められやすく、著作物として著作権法によって保護を受けられる可能性が高くなる。また、AIには(イ)を表現した者ではないため、原則としてAIが自律的に生成した表現は著作物として保護の対象にならない。

1. (ア)価値 (イ)知能
2. (ア)創造性 (イ)思想・感情
3. (ア)有用性 (イ)汎用性
4. (ア)精度 (イ)自律性

#### 問65

機械学習においてデータを用いて訓練し終え、これを予測タスクに使用できるモデルの名称として、最も適切な選択肢を1つ選べ。

1. 転移学習モデル
2. 事前学習モデル
3. 学習済みモデル
4. 汎用性モデル

#### 問66

3社がそれぞれ、以下のとおり特許を取得しているとする。

- ・X社は「機能Aを有するプログラム」の特許を取得している
- ・Y社は「機能Aと機能Bを有するプログラム」の特許を取得している
- ・Z社は「機能Aと機能Cを有するプログラム」の特許を取得している

以下の記述のうち、最も不適切な選択肢を1つ選べ。

なお、ライセンス契約(通常実施権や専用実施権の付与の契約)は行っていないとする。

1. X社は「機能Aと機能Bと機能Cを有するプログラム」を、業として実施できない。
2. Y社は「機能Aと機能Bを有するプログラム」を、業として実施できない。
3. Y社は「機能Aと機能Bと機能Cを有するプログラム」を、業として実施できない。
4. Z社は「機能Bと機能Cを有するプログラム」を、業として実施できない。

### 問67

カメラを用いて取得した情報のうち、個人情報保護法上、単体として個人情報とみなされるデータとして、最も適切な選択肢を1つ選べ。

1. カメラに写った人物の数を示すカウントデータ
2. カメラ画像にモザイク処理を施し、特定の個人が識別できないように加工したデータ
3. 画像から機械的に推定した人物の性別または年代
4. 目・鼻・口などの位置関係を抽出し、数値化したような顔の特徴量データ

### 問68

機械学習において、データの次元の数が増えれば増えるほど、計算量が指数関数的に増えてしまう現象として、最も適切な選択肢を1つ選べ。

1. ラッセルのパラドックス
2. モラベックのパラドックス
3. ノーフリーランチ定理
4. 次元の呪い

### 問69

人工知能(AI)の技術の社会実装において、考慮すべき倫理と法律の課題に関して、最も適切な選択肢を1つ選べ。

1. プライバシー・バイ・デザインとは、技術を実装する段階でプライバシー侵害がないことを確実に保護するというコンセプトである。
2. ELSI とは、AI の技術が社会に広く普及するための技術的な課題を解決する研究分野である。
3. サービスを通じて蓄積した個人情報を含むデータを分析に用いる際に、サービスの利用者の行動特性を評価するプロファイリングは、倫理の観点から、当該利用者に通知してはならない。
4. センシティブ属性をモデル訓練の特徴量に用いる場合、社会に混乱をもたらすアルゴリズムバイアスに厳重な注意が必要である。

### 問70

LSTM (Long Short-Term Memory)の派生版である GRU で 1 つの基本ユニットに含まれるゲート数として、最も適切な選択肢を 1 つ選べ。

1. ゲートは 1 つ
2. ゲートは 2 つ
3. ゲートは 3 つ
4. ゲートは 4 つ

### 問71

データの利用条件に関する注意点として、最も適切な選択肢を 1 つ選べ。

1. 著作物に該当するデータを利用する場合は、必ず著作権者から許諾を得る必要がある。
2. データを利用する目的に変更が生じた場合、既に一度本人からの利用の同意を得ているデータについては、利用目的の変更点について再度本人の同意を直接得ることが必要である。
3. 個人情報が含まれるデータを必ず匿名加工情報になるように加工する。
4. 人種・病歴・犯罪歴などの要配慮個人情報を学習データとして使用したい場合、原則本人の同意が必要である。

### 問72

OpenAI が 2020 年に公開した言語モデル GPT-3 の特徴として、最も不適切な選択肢を 1 つ選べ。

1. 教師なし学習を行う。
2. Transformer を用いる。
3. 従来モデルに比べてパラメータの数を削減できた。
4. 文章の生成ができる。

### 問73

ニューラルネットワークの訓練における学習率に関する説明として、最も不適切な選択肢を 1 つ選べ。

1. ハイパーパラメータである。
2. 勾配降下法においてモデルの最適化がうまくいくためには、学習の前半で小さく設定し、後半で大きく設定することが推奨される。
3. 収束するまでにかかる時間は学習率の値に依存する。
4. 学習率の値で重みを更新する歩幅が決まる。

#### 問74

以下の文章を読み、(ア)(イ)に入る用語の組み合わせとして、最も適切な選択肢を1つ選べ。

ニューラルネットワークをデータの分類に用いる場合、出力層には(ア)の数だけニューロン(ノード)があって、出力値は(イ)を表します。

1. (ア)クラス (イ)クラス
2. (ア)データ (イ)確率
3. (ア)クラス (イ)確率
4. (ア)データ (イ)クラス

#### 問75

個人データの正しい利用に関して、最も不適切な選択肢を1つ選べ。

1. カメラ画像から形状認識技術等を用いて人の形を判別し、その数量を計測した通行人カウントデータは、個人情報にあたらない。
2. 機微情報を取得するために本人の同意を得ることが条件である。
3. データを利用する目的に変更が生じた場合、既に一度本人からの利用の同意を得ているデータについては、利用目的の変更点を公表することで本人の同意とみなすことができる。
4. オプトアウト制度を利用することにより、事前に本人の同意を得なくてもその個人データを第三者に提供できることがある。

#### 問76

以下の文章を読み、空欄( )に入る用語として、最も適切な選択肢を1つ選べ。

個人情報取扱事業者は、( )を利用することにより、事前に本人の同意を得なくてもその個人データを第三者に提供することが可能である。

1. 第三者提供承諾制度
2. オプトアウト制度
3. 仮名加工制度
4. コンセンサス制度

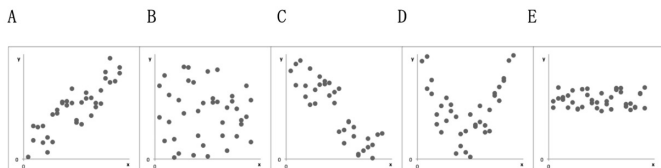
### 問77

深層強化学習において、ロボットの一連の動作を1つのディープニューラルネットワークによって実現する学習の仕組みとして、最も適切な選択肢を1つ選べ。

1. ノンストップ学習
2. 一気通貫学習(end-to-end 学習)
3. マルチモーダル学習
4. 敵対的学習

### 問78

下のグラフ A から E のうち、「正の相関」を示すもの、「負の相関」を示すものをそれぞれ全て選び、その組み合わせとして最も適切な選択肢を選べ。



1. 正 : A 負 : C
2. 正 : A, D 負 : C, D
3. 正 : A 負 : B, C, D
4. 正 : A, B, E 負 : C, D

### 問79

機械学習の文脈における GLUE の説明として、最も適切な選択肢を1つ選べ。

1. 汎化性能の弱いモデルを複数結合することで汎化性能の高いモデルを作る技術。
2. 自然言語理解(NLU ; Natural Language Understanding)タスクの精度を総合的に評価するためのベンチマーク。
3. より解釈性の高い自然言語処理モデルを実現するための研究分野。
4. CNN における畳み込み層とプーリング層の間に位置し、伝播されるデータを調整する関数。



### 問80

ニューラルネットワークの学習において、パラメータを一度にどの程度変更するかを表す量として、最も適切な選択肢を 1 つ選べ。

1. 歩幅
2. 更新幅
3. 加速度
4. 学習率

### 問81

AI モデルのエネルギー消費について、最も不適切な選択肢を 1 つ選べ。

1. パラメータの多いモデルほど一般的に計算量が多く、その学習で多くの電力が消費される。
2. 気候変動が進むにつれて、AI モデルの訓練による電力消費の問題がさらに深刻化すると思われる。
3. 自然言語処理は他の分野に比べてモデル訓練にかかる電力が少ない。
4. AI の省エネに向けた研究が発足している。

### 問82

過去のデータを用いて現在のデータを回帰で求めることで、時系列データの予測を行う分析モデルとして、最も適切な選択肢を 1 つ選べ。

1. ロジスティック回帰モデル
2. Lasso 回帰モデル
3. 自己回帰モデル
4. 単回帰モデル

### 問83

ルールが明確に決まっている簡単な問題にのみ対応でき、実世界の複雑な問題に対応できない AI 分野の問題として、最も適切な選択肢を 1 つ選べ。

1. ゲーム問題
2. 迷路問題
3. トイ・プロブレム
4. ナイーブ・プロブレム

#### 問84

新サービスの導入・設計にあたって、プライバシー・バイ・デザインの観点から望ましい状況として、最も不適切な選択肢を1つ選べ。

1. システムの利便性とプライバシーの保護の両方のバランスが考慮されている。
2. 企画・設計にすでにプライバシーを保護する対策を組み込まれている。
3. ユーザーのプライバシーが保護される仕組みが公に知られないような状態を担保した。
4. ユーザーによる設定を必要とせず、システムの初期設定においてユーザーのプライバシーが自動的に保護される。

#### 問85

標本抽出に関する記述として、最も不適切な選択肢を1つ選べ。

1. 全数調査が困難な場合、標本を用いて母集団の性質を調べる。
2. 無作為標本は、母集団と標本が「確率」を介して結びつけられる条件である。
3. 復元抽出とは「集団から取り出しては戻し、また取り出す」を繰り返す手法である。
4. 完全に無作為な「単純無作為抽出法」が一番楽な手法である。

#### 問86

空欄(ア)(イ)に入る組み合わせとして、最も適切な選択肢を1つ選べ。

MLOps とは、モデルを(ア)する側と(イ)する側が互いに連携しながら(ア)から(イ)までの一連のプロセスを管理する体制を指します。MLOps の各プロセスを繰り返すことによって、安定したシステムを維持することが重要である。

1. (ア)開発 (イ)運用
2. (ア)設計 (イ)開発
3. (ア)開発 (イ)改善
4. (ア)販売 (イ)評価

#### 問87

以下の文章を読み、空欄( )に入るモデルとして、最も適切な選択肢を1つ選べ。

Word2Vec の考案者でもあるトマス・ミコロフが発表した( )は、単語の分散表現を高速かつ高精度に学習でき、学習データに存在しない単語にも対応できることが特徴である。

1. fastText
2. Faster R-CNN
3. DCGAN
4. SHAP

### 問88

画像認識の CNN モデルを、スマホなどの小型な端末に使えるように工夫された内容として、最も適切な選択肢を 1 つ選べ。

1. MobileNet では、畳み込みを Depthwise（空間方向）と Pointwise（チャンネル方向）に分割して計算量を削減した。
2. MobileNet では、ネットワークの深さ、幅、解像度の最適化を Compound Coefficient（複合係数）という 1 つパラメータで調整した。
3. SegNet では、間隔をあけて畳み込み演算を行う Dilated 畳み込みが特徴である。
4. SegNet では、サイズの異なる複数の畳み込み層を並列に適用し、最後に連結する Inception モジュールが特徴である。

### 問89

データ活用プロジェクトを開始する前に設定した目標として、最もふさわしい選択肢を 1 つ選べ。

1. 売り上げデータを分析することで、販促を強化したい。
2. 物流データを活用して在庫の管理を最適化したい。
3. 販売個数を前期に比べて 30% 増やしたい。
4. 人事データを分析することで社員の離職率を食い止めたい。

### 問90

新しい画像を生成できるディープラーニングを用いたモデルとして、最も不適切な選択肢を 1 つ選べ。

1. FCN
2. VAE
3. Pix2Pix
4. DeepDream

### 問91

ニューラルネットワークのアーキテクチャ自体を最適化する技術として、最も適切な選択肢を 1 つ選べ。

1. RNN
2. GAN
3. NAS
4. DQN

## 問92

X 社が「機能 A と機能 B を有するプログラム」の特許を取得している。この状況についていえることとして、最も適切な選択肢を 1 つ選べ。

1. Y 社が「機能 A を有するプログラム(機能 B は有しない)」を開発したとき、このプログラムの販売は X 社の特許を侵害する。
2. Y 社が「機能 A と機能 B と機能 C を有するプログラム」を開発したとき、このプログラムの販売は X 社の特許権を侵害する。
3. Y 社が「機能 A と機能 C を有するプログラム(機能 B は有しない)」を開発したとき、このプログラムの販売は X 社の特許権を侵害する。
4. Y 社が「機能 A と機能 B と機能 C を有するプログラム」を開発したとき、このプログラムは新規性がなく、Y 社は特許を取得できない。

## 問93

ディープラーニングの活用に関して、最も適切な選択肢を 1 つ選べ。

1. ディープラーニングを用いた予測の根拠を解釈することが不可能であることは社会実装へのハードルとなる。
2. 構造化データにも使いづらい一方、非構造化データを用いた予測タスクでは、他の機械学習の手法よりも高い精度を出すことが多い。
3. ディープラーニングの手法を用いると特徴量を自動抽出できるため、他の機械学習よりもモデルのパラメータが少なく、学習が効率よく進む。
4. ディープラーニングは、識別タスクのみならず、データ生成タスクにも応用可能である。

## 問94

AI 開発やデータ分析では、Python 言語でサポートされているライブラリを用いることがある。このようなライブラリまたはフレームワークについて、最も不適切な選択肢を 1 つ選べ。

1. ライブラリを利用するためには中身を理解することが必須ではない。
2. TensorFlow はディープラーニングの実装に用いられる。
3. ライブラリを用いることで、比較的シンプルなコードで高度な機能を実装できる。
4. ライブラリは便利ではあるが、利用費が高価な場合が多い。

### 問95

以下の文章を読み、(ア)(イ)(ウ)の組み合わせとして最も適切な選択肢を1つ選べ。  
(ア)は今でも未解決であるAI分野の問題の1つである。この問題の本質は、解決しようとしている問題に関係する事象だけを取り出そうとする際に、検討しなければならない対象物の組み合わせが無数にあり、現実的な時間内で対応できない、ということである。(ア)は(イ)の収束する頃に明らかになった。実際現在使われている(ウ)は、特定の枠組みの中でのみ問題を解決することによって、この問題を疑似的に解決、もしくは回避しようとしている。

1. (ア)トイ・プロブレム (イ)第1次AIブーム (ウ)トイAI
2. (ア)フレーム問題 (イ)第2次AIブーム (ウ)汎用型AI
3. (ア)トイ・プロブレム (イ)第1次AIブーム (ウ)特化型AI
4. (ア)フレーム問題 (イ)第2次AIブーム (ウ)弱いAI

### 問96

強化学習における割引率の説明として、最も適切な選択肢を1つ選べ。

1. 最適化において、パラメータを一度にどの程度変化させるかを表す。
2. 最適化の各回において、学習データのうちどれくらいの割合を用いるかを表す。
3. どれくらいの確率でランダムな行動をとって新たな手を探索するかを表す。
4. 将来の報酬をどの程度現在の価値として考慮するかを表す。

### 問97

オートエンコーダの説明として、最も適切な選択肢を1つ選べ。

1. 隠れ層のノード数を入力層のノード数より多くすることにより、学習データのより詳細な特徴を抽出する。
2. 敵対的生成モデルにおいて、学習データの本物の画像か、ジェネレータの生成した画像かを判別する。
3. 深層信念ネットワークの構成要素である。
4. 入力層と出力層の値が一致するように学習する。

#### 問98

モデルの学習データに対して適用されるデータ拡張について、最も適切な選択肢を1つ選べ。

1. データ拡張を利用する場合は、増やした学習データの分を保存するためのコンピュータのメモリが十分であることを確認すべき。
2. 学習データにおけるクラスが不均衡である場合、数が少ないクラスを増やしバイアスを解消する方法である。
3. 学習データの1件1件を同じ倍数で数を増やす方法である。
4. データ拡張によって、データのバリエーションが増える。

#### 問99

指定した条件を満たすような画像を生成するニューラルネットワークとして、最も適切な選択肢を1つ選べ。

1. GAP
2. CGAN
3. Generator
4. DeepLab

#### 問100

学習データの収集、利用、第三者提供に際して、個人情報を含むデータを匿名加工情報になるように加工することがある。匿名加工情報に関する説明として、最も適切な選択肢を1つ選べ。

1. 個人情報を、特定の個人を識別できないように加工することで匿名加工情報になる。
2. 個人情報を含むデータを匿名加工情報に加工した後でも、本人の同意なく第三者提供することが禁じられている。
3. 匿名加工情報を作成する事業者には、匿名加工情報に関する苦情の処理や適正な安全管理措置の義務がある。
4. 匿名加工情報は仮名加工情報よりも、個人情報の保護レベルが低い。

### 問101

Google が開発した画像認識モデル DeepLab について述べた文章として、最も適切な選択肢を 1 つ選べ。

1. Atrous convolution を採用している。
2. 画像中の物体が存在する領域を矩形で切り出すモデルである。
3. U 字型の模式図で表すことのできる、対称的な構造を持つモデルである。
4. 物体のクラスの判定にサポートベクトルマシンが用いられている。

### 問102

画像認識モデルにおける Cutout の説明として、最も適切な選択肢を 1 つ選べ。

1. ネットワークのうち重要度の低いノードを取り除くことで、モデルを軽量化すること。
2. 画像中のランダムな位置を矩形領域で隠し、その状態でも予測できるように訓練することで、モデルの汎用性を高めるデータ拡張の一種である。
3. モデルの学習時に、ランダムに選んだノードを不活性化させることで、過学習を防ぐこと。
4. 学習データを訓練データとテストデータに分割し、訓練データのみで学習し、テストデータで汎化性能を評価すること。

### 問103

変分オートエンコーダ (VAE) の仕組みまたは特徴として、最も適切な選択肢を 1 つ選べ。

1. 従来の RNN に代わって、ゲート構造を設けることで長い系列データに対する予測精度が改善した。
2. 入力画像の特徴量を統計分布に変換し、その統計量を学習した結果をもとに新しいデータを生成する。
3. エンコーダ部分に CNN、デコーダ部分に RNN を持つ生成モデルである。
4. 被写体の些細な位置変化によって認識結果が変わらないように、入力画像データにおける重要な特徴を残しつつ情報量を圧縮するモデルである。

#### 問104

以下の文章を読み、空欄に最もよくあてはまる選択肢を1つ選べ。

複雑な構造を有するディープラーニングも、その計算処理の1つ1つは、単純な積和計算である。よって、大規模なテンソル計算が得意であり、膨大な単純作業の並列処理が得意で、従来グラフィック用途に利用されてきた( )がディープラーニング向けの処理に応用されている。

1. CPU
2. GPU
3. TPU
4. BPU

#### 問105

深層強化学習のモデルとして、最も不適切な選択肢を1つ選べ。

1. DQN (Deep Q-Network)
2. A3C
3. RAINBOW
4. LDA

#### 問106

ニューラルネットワークの基本構造であるパーセプトロンに関する説明として、最も不適切な選択肢を1つ選べ。

1. 単純パーセプトロンの中間層はユニットが少なく、構造がシンプルである。
2. 単純パーセプトロンを用いて非線形分離を行うことができる。
3. 多層パーセプトロンは、ニューロン層を必ず3つ以上を持つ。
4. 多層パーセプトロンは、必ず隠れ層を持つ。

#### 問107

ディープラーニングにおける量子化の説明として、最も適切な選択肢を1つ選べ。

1. 各層でデータを正規化することで、過学習を抑制すること。
2. パラメータのビット数を制限して、モデルを軽量化すること。
3. 重要度の低いノードを取り除き、パラメータを減らすこと。
4. 画像データに平行移動や回転などの変換操作を加え、学習データを水増しすること。



### 問108

以下の文章を読み、空欄( )に入る用語として、最も適切な選択肢を1つ選べ。

多層パーセプトロンとは、入力層、隠れ層、出力層の3種類のニューロン層から構成されるニューラルネットワークの総称である。入力層から訓練データを受け取り、隠れ層で識別に効果的な( )を抽出し、出力層から予測結果を出力する。

1. クラス
2. 画像
3. 特徴量
4. 報酬

### 問109

「相互情報量」とは2つの確率変数の相互依存性を表す指標であり、以下の式で表される。相互情報量の説明として、最も適切な選択肢を1つ選べ。

$$I(X; Y) = \int_Y \int_X p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy$$

1. 相互情報量は常に正の値をとる。
2. 2つの確率変数(XとY)を入れ替えても、相互情報量は同じ値をとる。
3. 2つの確率変数の相互依存性が小さいほど、相互情報量が大きくなる。
4. 相互情報量の式の中のlogは、指数関数を表す。

### 問110

機械学習モデルの汎化性能を高めるために、ハイパーパラメータの候補値を最適化する手法として、最も適切な選択肢を1つ選べ。

1. ブルーとフォース
2.  $\varepsilon$ -greedy 方策
3. ファインチューニング
4. ランダムサーチ

### 問111

AI プロジェクトを進めるための体制として、最も不適切なものを 1 つ選べ。

1. プロジェクトのチームには開発担当だけではなく、プロジェクトマネジャー、法務担当、広報も必要である。
2. システムの設計の段階から、UI や UX の使いやすさの観点からデザイナーが関わるべき。
3. 運用開始の時点から、システムの利用者のプライバシーの侵害を防ぐことを意識すべき。
4. AI システムを保守・運用するための人材を計画の段階で確保する必要がある。

### 問112

以下の文章を読み、空欄（ ）に入る用語として、最も適切な選択肢を 1 つ選べ。

強化学習アルゴリズム A3C において、3つの“A”は「Asynchronous」と「Advantage」と「Actor」を表し、“C”は「Critic」を表す。「Advantage」は「複数ステップ先を考慮してネットワークの重みを更新すること」を指し、「Asynchronous」は（ ）を指す。

1. 非連続にネットワークの重みを更新すること。
2. グローバルネットワークと重みを同期すること。
3. エージェントを非同期に学習させること。
4. 自律的な意思決定を学習させること。

### 問113

ビッグデータは「3つのV」で代表される特徴を持つ。この3つの特徴に含まれる要素として、最も不適切な選択肢を 1 つ選べ。

1. Volume（データ量が大きい）
2. Vague（データの漠然性が高い）
3. Velocity（データの生成と更新の速度が高い）
4. Variety（データに多様性がある）

### 問114

画像データからノイズを除去する操作として、最も適切な選択肢を 1 つ選べ。

1. トリミング
2. 平滑化
3. バッチ正規化
4. ブルーニング

### 問115

Mini-Max 法の説明として、最も適切な選択肢を 1 つ選べ。

1. 完全情報ゲームにおいて、自分の番で最もスコアの高い手を選び、相手の番で最もスコアの低い手を選ぶ戦略である。
2. 同じくゲーム理論の戦略である  $\alpha\beta$  法を、より効率よく計算できるように改良したものである。
3. 探索木において、一旦深さ方向に行けるところまで掘り下げてから、ゴールが見つからない場合に引き返して次の枝に移る探索戦略である。
4. ゲーム等のシミュレーションにおいて、ランダムに多数のパターンを確認することを通じて、最も適切な戦略を見つける手法である。

### 問116

以下の文章を読み、最も適切な用語の組み合わせの選択肢を 1 つ選べ。

ニューラルネットワークの学習の目的とは、(ア)の値が出来るだけ(イ)なるように、(ウ)の最適な値を見つけ、その値に設定を更新すること。

1. (ア)重み (イ)大きく (ウ)傾き
2. (ア)損失関数 (イ)小さく (ウ)パラメータ
3. (ア)誤差関数 (イ)大きく (ウ)隠れ層の数
4. (ア)活性化関数 (イ)小さく (ウ)傾き

### 問117

ビジネスにおけるデータ分析の留意点について述べたものとして、最も不適切な選択肢を 1 つ選べ。

1. データ分析のモデルを構築し終えた後に、分析モデルの運用に必要なタスクを明確化すべきである。
2. 目的とする分析が大量のデータを要する場合、事前にデータを収集、分析する環境を確保する。
3. まず小規模で PoC（コンセプト検証）を行い、目指す分析の精度、本番環境への適用可能性、諸リスクなどを把握することが推奨される。
4. 全て自動化に頼ることなく、データ環境の変化や予期せずリスクへ対処できるように、運用担当者が手動でシステムを調整する余地を残すべき。

#### 問118

以下の文章を読み、空欄( )に入る概念として、最も適切な選択肢を1つ選べ。ロボティクス分野にはディープラーニングが盛んに応用されている。特定の環境と状態における報酬を最大化するような一連の行動をロボットに学習させるために( )が使われることがある。

1. 半教師あり学習
2. 事前学習
3. 深層強化学習
4. ミニバッチ学習

#### 問119

機械学習におけるハイパーパラメータの特性として、最も適切な選択肢を1つ選べ。

1. ハイパーパラメータは、パラメータの中で一番予測精度への影響が高く、重要性が高いパラメータのこと。
2. ハイパーパラメータとは、学習のプロセスの中で自動的に最適化されるパラメータのこと。
3. ハイパーパラメータとは、学習済みモデルから他のモデルに転用できるパラメータのこと。
4. ハイパーパラメータを、モデルを構築する者が手動で設定できる値であり、その設定によってモデルの学習の状況が変わる。

#### 問120

以下の文章を読み、空欄( )に入る用語として、最も適切な選択肢を1つ選べ。ある時刻の状態がそれ以前の時刻の状態の影響を受けるようなデータは一般的に( )のことである。

1. GIS データ
2. 時系列データ
3. 非構造化データ
4. ユニークデータ

### 問121

RNN における教師強制 (Teacher Forcing) の説明として、最も適切な選択肢を 1 つ選べ。

1. 学習において、各時刻の入力に、1 つ前の時刻の出力データではなく、1 つ前の時刻の正解データを教師データとして用いる方法。
2. 学習において、各時刻の出力に、1 つ後の時刻の入力データを教師データとして用いる方法。
3. 学習において、各時刻の出力に、その時刻の入力データを教師データとして用いる方法。
4. 学習において、各時刻の入力に、その時刻の入力データではなく、1 つ前の時刻の出力データを教師データとして用いる方法。

### 問122

予測モデルの構築に用いられるデータにおけるバイアスについて、最も不適切な選択肢を 1 つ選べ。

1. 不均衡データを用いて分類モデルを学習する場合、学習済みモデルはデータ件数の少ないクラスを優先して損失関数を最小化する傾向にある。
2. ホールドアウト法で学習データを分割する際に、サンプリングバイアスが生じる可能性がある。
3. 複数の手法でモデルを構築しその結果を比べることでアルゴリズムバイアスに気づきやすくなる。
4. 偏りが精度に及ぼす影響はデータ全体のボリュームと関係する。

### 問123

2012 年に物体認識のコンペティション ILSVRC にて AlexNet モデルが優勝し、この出来事はディープラーニングが注目されたきっかけの 1 つとなった。ILSVRC で競技に参加したモデルの訓練に使用されたデータセットとして、最も適切な選択肢を 1 つ選べ。

1. MNIST
2. ImageNet
3. CIFAR-10
4. EfficientNet

#### 問124

マルチエージェント強化学習の説明として、最も適切な選択肢を 1 つ選べ。

1. 視覚情報(画像や動画)と聴覚情報(音声)など、複数種類のデータを用いた強化学習である。
2. 複数の主体が同時に協調して、あるいは、競争的に学習する。
3. ジェネレータとディスクリミネータの競争を通じて学習する。
4. 報酬の期待値が高いことが分かっている行動だけでなく、一定の確率でランダムな行動をとって新たな手を探索する。

#### 問125

ブロックチェーンの説明として、最も適切な選択肢を 1 つ選べ。

1. 強いセキュリティで管理された機関で管理されることで、データの改竄が困難である。
2. パスワードの設定がなくてもセキュリティを担保可能であることから、仮想通貨の用途に普及した。
3. デジタルの学習・資格証明書であるオープンバッジに活用されている。
4. ブロックチェーン上で保管される NFT アートは、作品そのものだけでなく著作権を含めて譲渡・売買されることで、資産価値が高まる。

#### 問126

K- 分割交差検証法に関する説明として、最も適切な選択肢を 1 つ選べ。

1. K 回の精度評価結果の平均値を、そのモデルの汎化性能と考える。
2. K 回の精度評価結果のうち、最も結果のよいものを、そのモデルの汎化性能と考える。
3. K 回の精度評価結果のうち、最も結果の悪いものを、そのモデルの汎化性能と考える。
4. K 回の精度評価結果のうち、最後に求めたものを、そのモデルの汎化性能と考える。

### 問127

混同行列を用いた機械学習の精度評価の説明として、最も適切な選択肢を1つ選べ。

1. 2つのモデルAとBを比較したとき、Bの適合率と再現率がともにAに比べて高いとき、Bの正解率もAに比べて高い。
2. 2つのモデルAとBを比較したとき、Bの適合率がAに比べて高いとき、Bの再現率はAに比べて低い。
3. 2つのモデルAとBを比較したとき、Bの適合率と再現率がともにAに比べて高いとき、BのF値もAに比べて高い。
4. 2つのモデルAとBを比較したとき、Bの正解率がAに比べて高いとき、Bの適合率と再現率の少なくとも一方はAに比べて高い。

### 問128

教師なし学習の使用例として、最も適切なものを1つ選べ。

1. 過去1年分の売り上げデータを学習データとし、来月の売り上げを予測するモデルを構築する。
2. ラベルづけられた猫と犬の画像を学習データに使い、新しい画像に対して、猫なのか犬なのかを分類できるモデルを訓練する。
3. 購買データ間の関連性を分析することで、ビールを買う人はナッツを買う確率が高いことがわかった。
4. 過去の医療診断事例を学習データとして使用し、年齢、性別、病歴、症状を特徴量として生存確率を予測するモデルを作る。

### 問129

文章のトピックを推定する際に用いられる技術として、最も不適切な選択肢を1つ選べ。

1. クラスタリング
2. 確率的潜在意味解析
3. チューリングテスト
4. 潜在的ディリクレ配分法

### 問130

音声認識において、音韻を区別する特徴となるピークが立っている周波数を表す用語として、最も適切な選択肢を1つ選べ。

1. ナイキスト周波数
2. フォルマント周波数
3. サンプリング周波数
4. クロック周波数

### 問131

深層学習のフレームワークでは、計算の流れをグラフ構造で描いた計算グラフにより、ネットワークの構造を視覚的に理解しながら設計することが多い。この計算グラフについて述べた記述として、最も適切な選択肢を1つ選べ。

1. Chainer が導入した define-by-run 方式は、順伝播処理を実行しながら計算グラフを動的に構築する。この方式は PyTorch でも使われている。
2. Chainer が導入した define-and-run 方式は、計算グラフを構築したのちに順伝播処理を実行する。この方式は PyTorch でも使われている。
3. Chainer が導入した define-and-run 方式は、順伝播処理を実行しながら計算グラフを動的に構築する。この方式は Caffe でも使われている。
4. Chainer が導入した define-by-run 方式は、計算グラフを構築したのちに順伝播処理を実行する。この方式は Caffe でも使われている。

### 問132

半教師あり学習を行う流れとして、最も適切な選択肢を1つ選べ。

1. 教師データがつく一部のデータで教師あり学習を行い、教師データがつかない一部のデータで教師なし学習を行い、両方の結果を平均化する。
2. 一部のデータのみ教師データがついており、これらを用いて教師あり学習を行い、残りのデータを使って精度をテストする。
3. 教師データがつく一部のデータで教師あり学習を行い、教師データがつかない一部のデータで教師なし学習を行い、両方の結果のうち精度の高い方を採用する。
4. 教師ラベルが付与されている一部のデータを利用し、残りの教師ラベルなしのデータに対してラベルの予測を行い、最後にすべてのデータを統合し教師あり学習を行う。



### 問133

空欄( )にあてはまる用語として、最も適切な選択肢を1つ選べ。

ニューラルネットワークの隠れ層では、( )がデータの値を整え、ネットワークの下の隠れ層で特徴を抽出しやすくする役割を持つ。

1. カーネル関数
2. デルタ関数
3. 活性化関数
4. 損失関数

### 問134

データを正規化する式は以下です。

$(\text{値} - \text{最小値}) / (\text{最大値} - \text{最小値})$

このような分子と分母を使う理由として、最も適切な選択肢を1つ選べ。

1. 外れ値の影響を取り除くため
2. 値を0と1の間におさめるため
3. 値を平均を中心に左右対象に分布させるため
4. ノイズをデータから除くため

### 問135

「機能Aを有するプログラム」の特許は存在するが、「機能Aと機能Bを有するプログラム」は公知ではなく、公然実施もされていない。このとき、「機能Aと機能Bを有するプログラム」の特許出願における新規性・進歩性について、最も不適切な選択肢を1つ選べ。

1. この出願は、新規性を有する。
2. この出願の技術が、当業者が想像できる程度を著しく超えて高い性能を示すとき、進歩性があると判断される。
3. この出願の技術が、産業上の利用可能性が高いとき、進歩性があると判断される。
4. この出願の技術が、当業者にとって「機能A」と「機能B」を組み合わせることが容易に思いつくものでないとき、進歩性があると判断される。

### 問136

一般物体認識の手法であるセマンティックセグメンテーションの特徴として、最も適切な選択肢を1つ選べ。

1. 画像の中に対象物が存在しそうな領域を矩形を用いて切り出して見つける。
2. 先に物体認識をし、続いてその物体を説明する自然言語を生成する総合モデル。
3. 画素の粒度で物体の輪郭を他の物体または背景から切り出すことができる。
4. 手書きの絵から写真風の画像を生成することができる。

### 問137

学習済みモデルの精度を検証するために、モデルの学習を始める前にあらかじめデータを「訓練データ」と「テストデータ」の2つに適切な割合で分割する必要がある。この割合として、最も適切な選択肢を1つ選べ。

1. 訓練：テスト = 3 : 7
2. 訓練：テスト = 7 : 3
3. 訓練：テスト = 2 : 8
4. 訓練：テスト = 5 : 5

### 問138

音声認識において周波数は音のどのような性質を表すか、最も適切な選択肢を1つ選べ。

1. 音の大小
2. 音の遠近
3. 音の高低
4. 音の強弱

### 問139

偽陽性の説明として、最も適切な選択肢を1つ選べ。

1. 正解が陽性であるサンプルを、機械学習モデルが陰性と判断してしまうこと。
2. 正解が陰性であるサンプルを、機械学習モデルが陽性と判断してしまうこと。
3. ある機械学習モデルで陽性と判断されたサンプルが、別の機械学習モデルでは陰性と判断されてしまうこと。
4. ある機械学習モデルで陰性と判断されたサンプルが、別の機械学習モデルでは陽性と判断されてしまうこと。

#### 問140

ビジネスにおいて、効率的にデータ活用・分析を行うために、ビジネス課題の理解、データの理解、データの準備、モデル作成、評価、展開・共有という6つのプロセスを規定したモデルとして、最も適切な選択肢を1つ選べ。

1. MLOps
2. DevOps
3. CRISP-DM
4. MECE

#### 問141

ディープラーニングの学習においてみられる学習につれて、テストデータに対する誤差が減少から増加に転じた後、再び減少に転じる現象を何というか、最も適切な選択肢を1つ選べ。

1. 二重降下現象
2. 鞍点
3. 早期終了
4. プラトー

#### 問142

Web APIの説明として、最も不適切な選択肢を1つ選べ。

1. HTMLの技術を用いて、ソフトウェア同士が機能の一部を共有することである。
2. 利用者は、APIのサーバーにリクエストを送り、それに対して応答が返ってくることで、プログラムの機能が利用できる。
3. APIの活用により、アプリケーションの開発時間の短縮やコスト削減につながる。
4. APIの活用により、自作でプログラムを組むよりも、より柔軟なアプリケーション開発が可能になる。

#### 問143

データリークageの説明として、最も不適切な選択肢を1つ選べ。

1. 本来予測時に得られないはずの情報がデータに含まれることである。
2. 学習時には高い精度が出たのに、本番環境では精度が出ないことにつながる。
3. 元のデータにはない摂動を加えたデータ汚染は、データリークageにあたる。
4. 時系列分析の学習において将来のデータが混ざってしまうのは、データリークageにあたる。

#### 問144

ミニバッチ学習の説明として、最も適切な選択肢を 1 つ選べ。

1. 最適化の各回において、データの一部をサンプリングして学習に用いること。
2. ネットワークのうち重要度の低いノードを取り除き、パラメータを減らすこと。
3. 学習済みの大規模なモデルの入出力の関係を、小規模なモデルに学習させること。
4. 全データを訓練データとテストデータに分割して、モデル構築すること。

#### 問145

強化学習モデルの構築において重要となることがある、将来の状態が現在の状態のみに依存し、過去の状態に依存しないという性質の名称として、最も適切な選択肢を 1 つ選べ。

1. 多重共線性
2. マルコフ性
3. 透明性
4. 限定提供性

#### 問146

自然言語処理モデルである Transformer の説明として、最も不適切な選択肢を 1 つ選べ。

1. 入力と出力の橋渡しに Source-Target Attention が用いられている。
2. エンコーダとデコーダには Self-Attention が用いられている。
3. 2 つの文が連続する文かどうかを判定する、Multi-Head Attention というタスクで事前学習されている。
4. Google の開発した自然言語処理モデル BERT に用いられている。

#### 問147

SegNet についての記述として、最も不適切な選択肢を 1 つ選べ。

1. セマンティックセグメンテーションのモデルである。
2. エンコーダとデコーダを有するモデルである。
3. 間隔をあけて畳み込み演算を行う Dilated 畳み込みを特徴とするモデルである。
4. バッチ正規化により学習の収束の高速化と安定化を実現させている。

#### 問148

大規模なデータセットで学習済みの畳み込みニューラルネットワーク(CNN)を用いて、転移学習により、自身の目的とするタスクに特化した画像識別モデルを作るとする。このとき、効率よくパラメータを最適化する手段として、最も適切な選択肢を1つ選べ。

1. 全結合層と、出力側に近い畳み込み層を最適化する。
2. すべての畳み込み層を最適化する。
3. 全結合層と、ランダムに選んだ畳み込み層を最適化する。
4. ランダムに選んだ畳み込み層を最適化する。

#### 問149

TD (Temporal Difference)学習の記述として、最も不適切な選択肢を1つ選べ。

1. 環境についての情報が使えないモデルフリーな強化学習として使える。
2. SALSA は TD 学習の一種である。
3. 報酬が得られるまで行動し、報酬を得た段階で過去の価値関数を一気に更新する。
4. TD 学習において Q 関数はベルマン方程式にもとづいて求める。

#### 問150

赤池情報量基準について述べた文章として、最も適切な選択肢を1つ選べ。

1. サポートベクトルマシンにおいて、誤分類を許容する基準である。
2. 重回帰分析において、最適な説明変数の組み合わせを選ぶ基準である。
3. 決定木において、どの軸を基準に枝分かれさせるかを選ぶ基準である。
4. 機械学習において、モデルの学習に必要なデータ量に関する経験的な基準である。

#### 問151

Seq2Seqの説明として、最も不適切な選択肢を1つ選べ。

1. エンコーダに CNN、デコーダに RNN を用いることで、画像の脚注生成が可能となる。
2. エンコーダに CNN、デコーダに CNN を用いることで、画像スタイルの変換が可能となる。
3. エンコーダに RNN、デコーダに RNN を用いることで、翻訳が可能となる。
4. エンコーダに RNN、デコーダに CNN を用いることで、画像分類が可能となる。

#### 問152

ディープラーニングにおいて、入力に重みをかけて、バイアスを加えてえられた値に適用する変換に関する用語として、最も適切な選択肢を 1 つ選べ。

1. アフィン変換
2. 活性化関数
3. フーリエ変換
4. スラック変数

#### 問153

深層強化学習のモデルである Rainbow の説明として、最も適切な選択肢を 1 つ選べ。

1. 2013 年に DeepMind 社が発表し、Atari 社のゲームで高いスコアを出して注目された。
2. ノイジーネットワーク、ダブル DQN、優先度付き経験再生など、当時新たな手法として知られた手法を組み合わせで高いスコアを出した。
3. 2016 年に世界的なトップ囲碁棋士に勝利したモデルであり、モンテカルロ木探索を組み合わせている。
4. 2011 年にアメリカのクイズ番組で歴代チャンピオンに勝利した。

#### 問154

ある二値分類モデルの適合率が 0.500、再現率が 0.750 であったとき、F 値として、最も適切な選択肢を 1 つ選べ。

1. 1.250
2. 0.612
3. 0.600
4. 0.625

#### 問155

EC サイトなどで顧客の購買を促す目的で、ユーザー同士の類似性およびユーザーの行動履歴に基づいて当該顧客の好みやニーズに合う商品を予測し推薦を行う手法として、最も適切な選択肢 1 つ選べ。

1. ユーザーフィルタリング
2. 行動フィルタリング
3. 協調フィルタリング
4. 類似フィルタリング

### 問156

重回帰分析において考慮すべき事項として、最も適切な選択肢を 1 つ選べ。

1. 説明変数の数を増やすほど、決定係数は小さい値をとる傾向がある。
2. 学習データ数は説明変数の数よりも多くなくてはいけない。
3. 説明変数の中に、互いに独立な変数同士が含まれてはいけない。
4. その現象を説明するのに、最低限必要な説明変数の数の目安として、赤池情報量基準(AIC)が知られている。

### 問157

以下の文章を読み、(ア)(イ)(ウ)にあてはまる組み合わせとして、最も適切な選択肢を 1 つ選べ。

ELSI は、生命科学分野の「ヒトゲノム計画」から生まれ、現在 AI・データサイエンスの分野でも重視されているという研究分野である。ELSI の E は(ア)、L は(イ)、I は Issues (課題)、S は(ウ)の略である。

1. (ア) Ethics (倫理)            (イ) Life (生命)            (ウ) Safety (安全)
2. (ア) Ethics (倫理)            (イ) Legal (法律)            (ウ) Society (社会)
3. (ア) Economy (経済)            (イ) Life (生命)            (ウ) Society (社会)
4. (ア) Economy (経済)            (イ) Legal (法律)            (ウ) Safety (安全)

### 問158

自然言語処理における「照応解析」の説明として、最も適切な選択肢を 1 つ選べ。

1. 代名詞などの表現が何を指しているのかを推定すること
2. 2 つの文が連続する文かどうかを判断すること
3. 文中における主語と述語の対応関係を推定すること
4. 文中の各単語がどの品詞に対応するかを推定すること

### 問159

単純パーセプトロンの問題点として、最も不適切な選択肢を 1 つ選べ。

1. 線形分離不可能な分類問題には対応できない。
2. XOR 回路を再現することができない。
3. 分類問題は、二値分類問題しか解けない。
4. 回帰問題には対応できない。

#### 問160

正則化には、パラメータの絶対値の和を罰則項として加えて特徴選択と次元圧縮を行うラッソ回帰や、パラメータの二乗和を罰則項として加えて滑らかな過学習を防止するリッジ回帰があるが、その両方を組み合わせた手法の名称として、最も適切な選択肢を1つ選べ。

1. Elastic Net
2. EfficientNet
3. GoogLeNet
4. AlexNet

#### 問161

CNNで行われる畳み込み演算の計算処理について考える。 $5 \times 5$ のサイズの画像に対して、パディングを1とし、 $3 \times 3$ のフィルタをストライド1で適当した場合の出力の図のサイズを答えよ。

1.  $5 \times 5$
2.  $7 \times 7$
3.  $3 \times 3$
4.  $6 \times 6$

#### 問162

自然言語処理に用いられる Attention の説明としてとして、最も不適切な選択肢を1つ選べ。

1. 文章中の単語間の関連度を計算することで、各単語にどれくらい注目すべきかを学習することができる。
2. 入力文内または出力文内の単語間の関連度を計算したものを Self-Attention という。
3. RNN モデルと比較して、離れた位置にある単語同士の関係性が捉えやすくなるが、その分、並列処理ができなくなる。
4. RNN モデルと比較して、文脈を考慮したより適切な分散表現を得ることができる。



### 問163

音声認識におけるメル尺度についての記述として、最も適切な選択肢を 1 つ選べ。

1. アナログな音声データからサンプリング定理に基づいて標準化する際の標本間隔
2. 音色の違いを表すスペクトル上の穏やかな変動を表す係数列
3. 音韻を区別する特徴となるピークが立っている複数の周波数
4. 人間が知覚する音の高さを表現するため、周波数を非線形変換した尺度

### 問164

機械学習モデルの開発を自動化する仕組みを表す用語として、最も適切な選択肢を 1 つ選べ。

1. AutoML
2. RPA
3. モンテカル口法
4. エキスパートシステム

### 問165

最小二乗法の説明として、最も適切な選択肢を 1 つ選べ。

1. 二乗誤差を最大化するパラメータを求める。
2. 二乗誤差を最小化するパラメータを求める。
3. 絶対誤差を最大化するパラメータを求める。
4. 絶対誤差を最小化するパラメータを求める。

### 問166

画像データベース ImageNet のクラス名の参照元となった、上位語 / 下位語の概念を取り入れた概念辞書の名称として、最も適切な選択肢を 1 つ選べ。

1. Word2Vec
2. Attention
3. WordNet
4. AlexNet

#### 問167

個人情報の保護に関する法律(個人情報保護法)の説明として、最も適切な選択肢を1つ選べ。

1. 外国に居住する外国人の個人情報は、個人情報保護法の保護の対象ではない。
2. 個人情報を加工することで匿名加工情報に変換した後も、扱い上個人を特定できる情報と照合しないように注意する必要がある。
3. 個人情報保護法における個人情報取扱事業者は営利企業に限定される。
4. メールアドレスのユーザー名及びドメイン名を用いて、特定の個人を識別可能である場合、メールアドレス単体が個人情報となる。

#### 問168

DCGAN について述べた文章として、最も適切な選択肢を1つ選べ。

1. 自然言語処理を行うエンコーダとデコーダを有するモデルである。
2. 画像生成を行うジェネレータとディスクリミネータを有するモデルである。
3. エンコーダとデコーダを有する強化学習モデルである。
4. ジェネレータとディスクリミネータを有する音声認識モデルである。

#### 問169

NAS (Neural Architecture Search)の説明として、最も不適切な選択肢を1つ選べ。

1. NAS の計算量を削減し、探索の効率化した技術の例として、ResNet が挙げられる。
2. ニューラルネットワークの構造を自動的に最適化できる技術である。
3. 一般に、NAS は、大量の GPU が必要で、時間もかかることが多い。
4. 強化学習の考え方をを用いて、ニューラルネットワークの構造自体を探索する。

#### 問170

個人情報保護法にもとづく個人情報を含むデータについて、最も適切な選択肢を1つ選べ。

1. 株式会社に関する情報も個人情報に該当する可能性がある。
2. 匿名加工情報とは、それだけで特定の個人を識別することができないように個人情報を加工したものと定義される。
3. 顧客との電話の通話内容を録音したログデータは、通話内容から特定の個人を識別できない場合でも、個人情報に該当することがある。
4. 要配慮個人情報を例外とし、個人データに対してマスキングを行えば、匿名加工情報になる。

#### 問171

オフライン強化学習の説明として、最も適切な選択肢を1つ選べ。

1. 環境に関するパラメータを明示的に推定せずに方策を求める強化学習であり、環境のモデルを推定することは困難である場合に用いられる。
2. 複数のエージェントによる非同期な並列学習を行うことを特徴とする強化学習であり、経験の自己相関を低減することで学習を安定化させることができる。
3. 事前に集めたデータを用いる強化学習であり、センサーログなど、大量のデータが事前に入手に可能な場合に用いられる。
4. シミュレーションではなく実環境と相互作用しながら学習を行う強化学習であり、パフォーマンスギャップを低減させることができる。

#### 問172

活字や手書き文字の画像を認識し、デジタルな文字データに変換する技術やソフトウェアを表す用語として、最も適切な選択肢を1つ選べ。

1. DTP
2. RPA
3. PCA
4. OCR

#### 問173

画像分類のニューラルネットワークに用いられるグローバルアベレージプーリング(GAP)の説明として、最も適切な選択肢を1つ選べ。

1. 畳み込み層の最後でカテゴリ分類のために用いられる処理である。
2. 2012年のILSVRCで優勝したAlexNetで初めて採用されたプーリング処理である。
3. 特徴マップの局所ごとに値の平均値を求める。
4. 学習データにバリエーションを持たせるためのデータ拡張の技術である。

#### 問174

ある事象が一定の時間内に発生する回数を表す確率分布として、最も適切な選択肢を1つ選べ。

1. 正規分布
2. t分布
3. ポアソン分布
4. 二項分布

### 問175

2014 年の ILSVRC で優勝した ResNet の特徴として、最も不適切な選択肢を 1 つ選べ。

1. 誤差率が人間の分類精度を初めて超えたといわれている。
2. 畳み込み層の一部を複数スキップする構造を導入して、勾配消失問題を軽減した。
3. 残差ブロックを用いて勾配を奥深い入力層側まで伝達させることで、高精度化した。
4. ネットワークの深さ、幅、解像度の最適化を Compound Coefficient（複合係数）という 1 つパラメータで調整した。

### 問176

以下の文章を読み、(ア)～(イ)の組み合わせとして最も適切な選択肢を 1 つ選べ。  
EU 一般データ保護規則(GDPR)の特色の 1 つとして(ア)の権利を認めている。これはあるサービスに対して、そのユーザーが自らに関して収集・蓄積した利用履歴などのデータを、ほかのサービスでも利用可能な形で移転可能にすることを求める権利である。欧州委員会によると、この権利は個人データにおいてユーザーの管理権限の強化とともに(イ)という意義がある。

- |                     |                       |
|---------------------|-----------------------|
| 1. (ア)データトランスペアレンシー | (イ)自動処理に基づく意思決定の透明性   |
| 2. (ア)データポータビリティ    | (イ)新興企業による新規サービス創出を促す |
| 3. (ア)データトランスペアレンシー | (イ)新興企業による新規サービス創出を促す |
| 4. (ア)データポータビリティ    | (イ)自動処理に基づく意思決定の透明性   |

### 問177

音声認識において、音声信号を周波数スペクトルに変換する操作として、最も適切な選択肢を 1 つ選べ。

1. フーリエ変換
2. ロジット変換
3. アフィン変換
4. モダリティ変換

### 問178

DeepMind 社が 2016 年に発表した音声合成のニューラルネットワークのモデルであり、従来のモデルに比べて自然な音声を作成することに成功したモデルの名称として、最も適切な選択肢を 1 つ選べ。

1. EfficientNet
2. SegNet
3. WaveNet
4. AlexNet

### 問179

以下の文章を読み、空欄(ア)～(ウ)に最もよくあてはまる選択肢を 1 つ選べ。  
機械学習の制度評価において、分類の閾値を変えた場合に偽陽性率と真陽性率がどのような関係になるかをプロットした図を(ア)といい、グラフの曲線より下の部分の面積を(イ)という。モデルがデータの特徴を全く捉えられずランダムに分類しているとき、(イ)の値は(ウ)に近い値となる。

1. (ア) ROC (イ) AUC (ウ) 0.5
2. (ア) AUC (イ) ROC (ウ) 0.5
3. (ア) ROC (イ) AUC (ウ) 0
4. (ア) AUC (イ) ROC (ウ) 0

### 問180

サードパーティークッキーに関する記述として、最も不適切な選択肢を 1 つ選べ。

1. クッキーを個人関連情報とみなし厳しい規制が導入される国や地域が増えている。
2. Chrome や Safari などの、ブラウザ側でも規制を進めている。
3. 閲覧しているサイトのドメインが発行するクッキーのことである。
4. ユーザー側はそのクッキーの発行元に情報を収集されていることに気づかないことが多い。

#### 問181

リカレントニューラルネットワーク(RNN)に特徴的な再帰構造の効果について述べた記述として、最も適切な選択肢を1つ選べ。

1. 勾配消失問題を抑制する。
2. 過学習を抑制する。
3. 時系列データの分析を可能にする。
4. わずかな位置の平行移動に対してロバストになる。

#### 問182

ディープラーニングは、機械学習の中でも特にパラメータの数が多いことから、どのような特徴を有するか、最も適切な選択肢を1つ選べ。

1. より少ない学習データ数で、目的に合った特徴量を抽出できる。
2. モデルが柔軟になり、過学習が起こりにくくなる。
3. 出力結果がより人間にとって納得感の高い、解釈性の高いモデルとなる。
4. パラメータの最適化に要する時間が長くなる。

#### 問183

重回帰分析によるモデル構築が可能な事例として、最も適切な選択肢を1つ選べ。

1. より長い距離を安定して歩行できるロボットを訓練する。
2. 入力された画像を説明する文章を作成する。
3. チェーン店の来客数を、その地域人口、店舗の面積、店舗の従業員数から予測する。
4. メールを、スパムメールとスパムメールではないメールに分類する。

#### 問184

音声認識において、音色の違いを表すスペクトル上の穏やかな変動(共振特性)を表す用語として、最も適切な選択肢を1つ選べ。

1. ケプストラム
2. メル周波数
3. スペクトル包絡
4. レゾナンス

### 問185

勾配爆発問題に関する説明として、最も不適切な選択肢を 1 つ選べ。

1. 勾配爆発問題への対策として、勾配が一定の値以上にならないようにする勾配クリッピングという方法が知られている。
2. 勾配爆発がおこると学習が不安定化する。
3. 勾配爆発は、RNN に特徴的な再帰的構造により、時間方向に積み重なった多層構造となっていることが原因の 1 つである。
4. 勾配爆発は、長期の時間依存性を考慮できない問題を解決するために、ゲート構造を設けたことが原因の 1 つである。

### 問186

サポートベクトルマシンにおけるマージンの説明として、最も適切な選択肢を 1 つ選べ。

1. 非線形な分類を可能にするための方法。
2. 分類の境界と、境界に最も近いデータの間の距離。
3. 誤分類をどれくらい許容するかの指標。
4. データの次元数(特徴量の数)。

### 問187

画像認識では CNN を用いるのが一般的である理由として、最も適切な選択肢を 1 つ選べ。

1. 画像データは通常、行列の大部分が 0 となるスパースなデータであるため。
2. 縦横の位置関係が重要であり、元の画像の次元を維持したまま計算を進めることが必要だから。
3. 時間依存性を考慮するために、再帰的なネットワーク構造が必要であるから。
4. 十分量の学習データを事前に集めることが難しく、現実の環境との相互作用に基づいてモデルを最適化することが重要であるから。

### 問188

学習率の設定におけるモーメンタムの説明として、最も不適切な選択肢を 1 つ選べ。

1. 現在の勾配だけでなく、過去の勾配を取り入れ、“慣性”の考え方を導入している。
2. Adam 法は、モーメンタムの考え方を取り入れて発展させたものである。
3. 一定の確率  $\epsilon$  でランダムな動きをすることで、局所最適解を回避することができる。
4. 谷や鞍点や局所最適解付近での振動を抑制することができる。

### 問189

自動運転の位置推定などに使われ、1 時刻前の状態の推定値と、現在についての複数の種類の測定値に基づいて、現在の状態の推定値を求める状態空間モデルとして、最も適切な選択肢を 1 つ選べ。

1. Rainbow
2. A3C
3. ARIMA
4. カルマンフィルター

### 問190

主成分分析に関する記述として、最も適切な選択肢を 1 つ選べ。

1. 相関を持つ特徴量を組み合わせ、分散が最小になるように次元を圧縮する。
2. 確率分布として  $t$  分布を仮定することで、複雑で高次元の構造を次元圧縮できる。
3. 線形な次元圧縮手法であり、非線形データには適用できない。
4. 圧縮後の次元は、2 次元または 3 次元である。

### 問191

誤差逆伝播法を用いて学習するニューラルネットワークに用いる活性化関数に求められる条件として、最も適切な選択肢を 1 つ選べ。

1. 微分可能である必要がある。
2. 値域が 0 から 1 の間である必要がある。
3. 定義域の全区間で積分すると 1 になる必要がある。
4. 線形関数である必要がある。

### 問192

ディープラーニングの学習では、局所最適解に陥りにくくし、異常値の影響を抑える目的で、ランダム性を取り入れた確率的勾配降下法がよく用いられる。このランダム性の説明として、最も適切な選択肢を 1 つ選べ。

1. 勾配情報にランダムなノイズを加えて、出力層側から入力層側へと伝播させる。
2. 勾配情報を用いてパラメータ更新をする際の学習率をランダムに選ぶ。
3. 学習データの中から、その都度ランダムに選んだサンプルを用いて学習を行う。
4. 一定割合のノードをランダムに不活性化させながら学習を行う。