

情 報

1

次の生成 AI（Artificial Intelligence；人工知能）に関する説明文の空欄 ～
 に入る最も適切な言葉を選択肢から選び、それぞれ番号をマークしなさい。

2023年 9 月 7 日に開催された広島 AI プロセス閣僚級会合において、生成 AI を含む高度な AI システムについて、, , , などリスクが挙げられている。

について、現状の生成 AI を含め、AI システム全般として、AI システムへの入力および、プロンプト（生成 AI への入力）に対して、AI システムはなんらかの計算を行い、結果を導出し提示するわけであるが、現在の AI システムのほとんどにおいて、どうしてその結果が導出されたかが分からない になっているという問題がある。

について、特に 形式のようなリアルタイムの会話形式でコミュニケーションをとることができるシステムに生成 AI を導入した場合、ハルシネーションと呼ばれる嘘の結果を出力される場合がある。例えば、生成 AI に多く利用されている である Transformer は関連のある語やフレーズを集約して文章を生成するのみの機能であり、決して正しい知識を確認した結果ではないことに留意すべきである。また、画像、音声、動画を対象とした生成 AI で偽物の画像、音声、動画が現実の映像・音声に限りなく近い形で生成されることがある。上記のような が などで拡散され、最悪、社会をパニックに陥れることが危惧されている。

について、生成 AI によって作成された作品の 権が誰に帰属するかは未解決の問題である。また、生成 AI は既存のデータを学習して新しい作品を生成するため、訓練データに含まれる一部を無断で使用され、既存の 物の権利を侵害するリスクがある。また、そもそも学習に用いられる既存データがどのように収集され、利用されているかに関する が必要ではないかという議論がある。

について、生成 AI の運用中に収集される学習のためのデータが適切に保護されていない場合、データ漏洩や不正アクセスのリスクがある。特に生成 AI のプロンプトに を入力してしまうことにより、それが第三者に漏洩する可能性がある。

上記の問題があるが、生成 AI は人間の学習や仕事の効率を大幅に向上させるツールとなり得る。また、自身の発想支援にも用いることが可能であり、リテラシーを高め、正しく利活用することが重要である。

選択肢

- | | | | |
|------------|----------|---------|----------|
| ① 透明性 | ② 偽情報 | ③ 知的財産権 | ④ 著作 |
| ⑤ チャット | ⑥ アルゴリズム | ⑦ SNS | ⑧ プライバシー |
| ⑨ ブラックボックス | ⑩ 個人情報 | | |

2

下記の文章を読み、次の各問い（問 1 ～ 問 5）に答えなさい。

表 1 は、2024年にパリで開催されたオリンピックのメダルランキングと人口、GDP、一人当たりの GDP のデータをまとめたものである。

表 1

メダル ランキング	国・地域	メダル獲得数 (個)	人口 (百万人)	GDP (10億 US ドル)	国民一人当たりの GDP (US ドル)
1	アメリカ合衆国	126	334.24	25440	76113
2	中国	91	1411.75	17960	12722
3	日本	45	123.95	4232	34143
4	オーストラリア	53	26.67	1693	63480
5	フランス	64	67.60	2779	41109
6	オランダ	34	17.68	1009	57070
7	イギリス	65	67.33	3089	45879
8	韓国	32	51.73	1674	32360
9	イタリア	40	58.85	2050	34834
10	ドイツ	33	83.20	4082	49063
...	...	...	...	...	...
84	カタール	1	2.69	236.3	87661
84	シンガポール	1	5.63	466.8	70100
84	スロバキア	1	5.43	115.5	21271

表 1 からメダルランキング上位30カ国のデータを抽出し、メダル獲得数と、人口、GDP、一人当たりの GDP との関係を可視化するため散布図を作成した。また、回帰分析を行い、求めた回帰直線を散布図と合わせて描画した。(図 1 ～ 3)

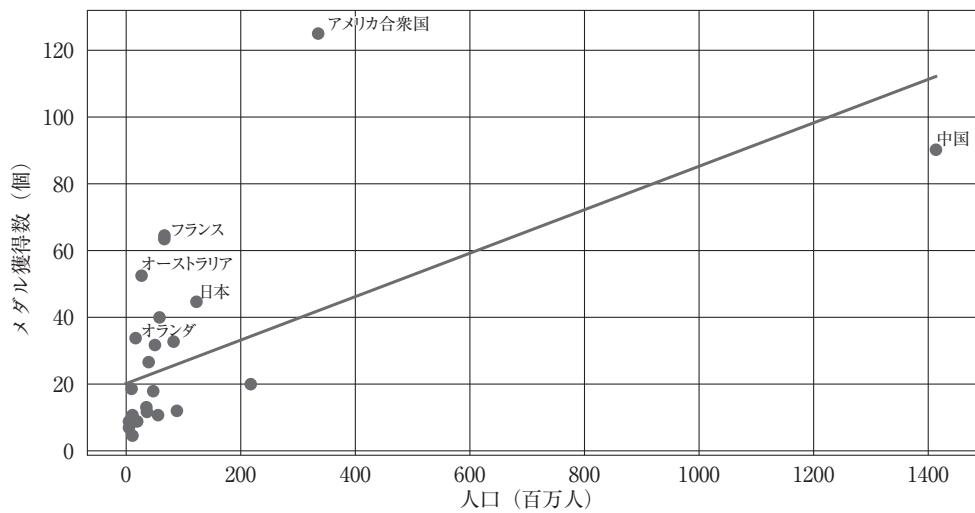


図1 人口とメダル獲得数 (メダルランキング上位30カ国)

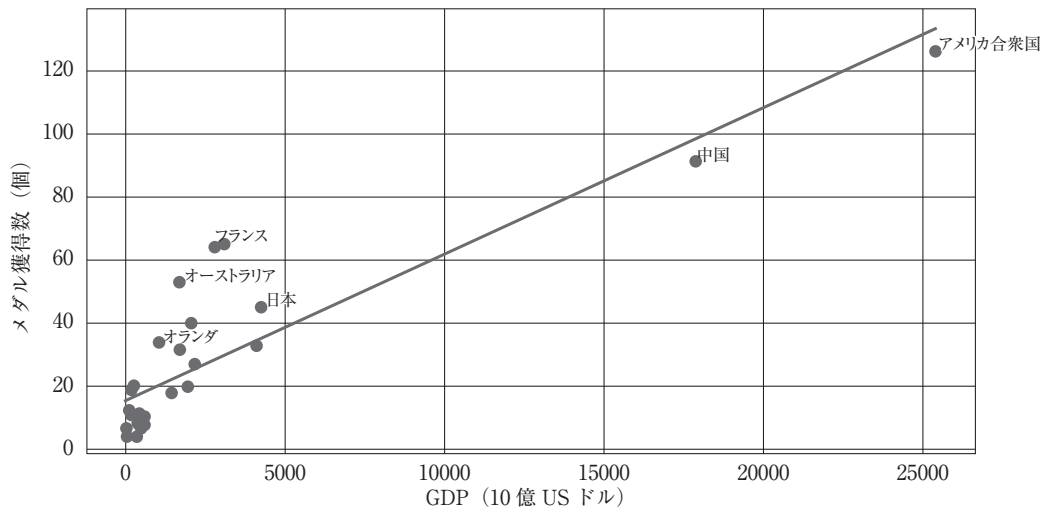


図2 GDP とメダル獲得数 (メダルランキング上位30カ国)

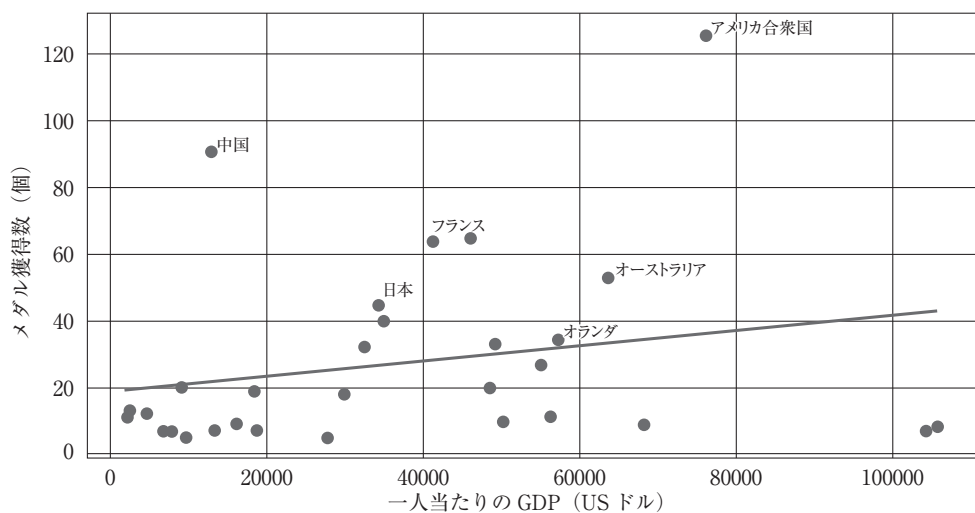


図3 一人当たりの GDP とメダル獲得数 (メダルランキング上位30カ国)

メダル獲得数の合計と人口，GDP，一人当たりのGDPとの相関係数をそれぞれ計算したところ次のようになった。

メダル獲得数と人口：0.59

メダル獲得数とGDP：

メダル獲得数と一人当たりのGDP：

メダル獲得数の合計と人口，GDP，一人当たりのGDPとの回帰式はそれぞれ次のようになった。

メダル獲得数と人口： $y = 0.065x + 21$

メダル獲得数とGDP： $y = \text{オ}x + \text{カ}$

メダル獲得数と一人当たりのGDP： $y = \text{キ}x + \text{ク}$

さらに，求めた回帰式の精度について検証するため，下記の式から決定係数 R^2 を計算した。決定係数 R^2 は，1に近ければ近いほど精度が高いといえる。

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \hat{y})^2}$$

y_i ：観測値

$\hat{y}_i$ ：回帰式による推定値

$\hat{y}$ ：観測値の平均

n ：データ数

メダル獲得数と人口： $R^2 = 0.35$

メダル獲得数とGDP： $R^2 = \text{ケ}$

メダル獲得数と一人当たりのGDP： $R^2 = \text{コ}$

問1 相関係数の特徴として、最も適切な文章を選択肢の中から選び、その番号をマークしなさい。

選択肢

- ① 相関係数が高いデータ間には因果関係があると考えられる
- ② データの数が半分になると相関係数も半分の値になる
- ③ 相関係数0.5と相関係数 -0.9 のデータでは、後者の方が相関が弱い
- ④ 相関係数は、その値の絶対値が大きい方が関連性が高いといえる

問2 とにあてはまるものを次の選択肢の中から選び、その番号をマークしなさい。

選択肢

- ① 0
- ② -0.24
- ③ 0.24
- ④ -0.59
- ⑤ 0.59
- ⑥ -0.88
- ⑦ 0.88

問3 回帰分析の説明として最も適切なものを選択肢の中から選び、その番号をマークしなさい。

選択肢

- ① 回帰分析は、データ間の因果関係の検証を目的としている
- ② 回帰分析は、データの分類を目的としている
- ③ 回帰分析は、データ間の関係をモデル化し、データの予測を目的としている
- ④ 回帰直線の傾きが正である場合、説明変数と目的変数には関係がないといえる

問4 , , , にあてはまるものを次の選択肢の中から選び、その番号をマークしなさい。

選択肢

- ① 0
- ② 15
- ③ 19
- ④ 25
- ⑤ 0.0002
- ⑥ 0.004
- ⑦ 0.06

問5 , にあてはまるものを次の選択肢の中から選び、その番号をマークしなさい。

選択肢

- ① -0.001
- ② 0.001
- ③ -0.35
- ④ 0.35
- ⑤ -0.77
- ⑥ 0.77

3

N 進法に関する下記の文章を読み、次の各問い（問 1 ～ 問 5）に答えなさい。

(1) 任意の N 進法において、正の整数を表示するデバイスを作ることを考える。以下では、1つの数字を表示するのに1つの部品が必要であると考え。例えば2進法ならば、0と1の2種類の数字を表示する部品を桁数分だけ組み合わせれば良い。それに対して10進法の場合は、0から9までの10種類を表示する部品を桁数分組み合わせることになる。

以下にデバイス A とデバイス B の2つのデバイスを示す。デバイス A は2進法のデバイスであり、デバイス B は10進法のデバイスである。ここでは例としてそれぞれのデバイスで、10進法における数値の12を表しており、表示に使われている部品を灰色で示している。なお、以下では、特に注記しない限り、数値は10進法の値として扱う。

デバイス A									
2^9	2^8	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1

図 1 2進法で数値を表示するデバイス

デバイス B	
10^1	10^0
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9

図 2 10進法で数値を表示するデバイス

問1 以下の文における、～に当てはまる数値をマークしなさい。

- ・デバイス A で表示できる数値の最大値は である。
- ・デバイス B で表示できる数値の最大値は99である。
- ・デバイス A の部品の数は である。
- ・デバイス B の部品の数は である。

(2) デバイス A とデバイス B で部品の数は同じである。それに対して、表示する事ができる数値の種類の数には違いがある。同じ部品の数でより多くの数字が表示できる場合、それはそのデバイスの方が効率が良いと考えることができる。

問2 以下の文における、(a)，(b) に当てはまる語の組み合わせとして適切なものを、選択肢から選んで に番号をマークしなさい。

- ・デバイス A とデバイス B を比べたとき、より多くの種類の数値を表せるのは (a) である。したがって、より効率が良いのは (b) である。

選択肢

- ① (a)：デバイス A，(b)：デバイス A
- ② (a)：デバイス A，(b)：デバイス B
- ③ (a)：デバイス B，(b)：デバイス A
- ④ (a)：デバイス B，(b)：デバイス B

(3) 以下の図3に示すのは2進法で数値を表すデバイスC，図4に示すのは3進法で数値を表示するデバイスD，図5に示すのは4進法で数値を表示するデバイスEである。
この3つのデバイスは部品数が同じで，どれも12個である。

デバイスC

2^5	2^4	2^3	2^2	2^1	2^0
0	0	0	0	0	0
1	1	1	1	1	1

図3 2進法で数値を表示するデバイス

デバイスD

3^3	3^2	3^1	3^0
0	0	0	0
1	1	1	1
2	2	2	2

図4 3進法で数値を表示するデバイス

デバイスE

4^2	4^1	4^0
0	0	0
1	1	1
2	2	2
3	3	3

図5 4進法で数値を表示するデバイス

問3 以下の文における，ク～スに当てはまる数値をマークしなさい。

- ・デバイスCで表示できる数値の種類はクケ種類である。
- ・デバイスDで表示できる数値の種類はコサ種類である。
- ・デバイスEで表示できる数値の種類はシス種類である。

問4 効率が良いことを，比較したときに値が大きいとして表現できると仮定する。デバイス C，デバイス D，デバイス E の3つのデバイスの効率の大小関係として適切なものを選択肢から選んで

セ

 に番号をマークしなさい。

選択肢

- ① デバイス C = デバイス D > デバイス E
- ② デバイス C > デバイス D = デバイス E
- ③ デバイス C > デバイス D > デバイス E
- ④ デバイス D > デバイス C = デバイス E
- ⑤ デバイス D > デバイス E > デバイス C
- ⑥ デバイス D = デバイス E > デバイス C
- ⑦ デバイス E > デバイス C = デバイス D
- ⑧ デバイス E > デバイス C > デバイス D
- ⑨ デバイス E > デバイス D > デバイス C

(4) 任意の正の整数を与えられたときに、0 からその数値までのすべての整数を表現することができるデバイスを作るとして、何進法で作ったら部品が少なくて済むのかを考える。このとき、桁数は数値を表現するのに最小限で構成することとする。また、桁を追加する際には、その桁で表現できる数値の種類の数だけの部品を用意しなくてはならないとする。例えば、3 進法でデバイスを作るとして、新たな桁を追加する際には、0, 1, 2 を表す 3 つの部品を追加しなくてはならず、0, 1 の 2 種類だけを追加することは出来ない。

問 5 以下のそれぞれの数値について、0 からその数値までのすべての整数を表現するデバイスを作るとして、選択肢の中から最も部品が少なくなるものをそれぞれ選んで ～ に当てはまる番号をマークしなさい。同じ選択肢を繰り返し選んでもよい。

- ・ 127 :
- ・ 240 :
- ・ 24 :

選択肢

- ① 2 進法
- ② 3 進法
- ③ 5 進法
- ④ 7 進法

4

山田さんは、入力としてじゃんけんのグー、チョキ、パーを与え、その入力について必ず勝つような出力を導出するアルゴリズムを考えた。下記の文章を読み、次の各問い（問1～問4）に答えなさい。

まずは、入力のデータ構造を考えた。⁽¹⁾ワンホットベクトル表現手法を用いて値が0か1を取るベクトルを考えた。具体的には、図1のように、グー、チョキ、パーのため、入力を3つの値からなる列ベクトル（3つの値を縦に並べたもの）を設定する場合、どれか1つの値が1でその他の値が0であるならば、グー、チョキ、パーを一意に特定することが可能である。

グー	チョキ	パー
1	イ	0
0	ウ	0
0	エ	1

図1 グー、チョキ、パーを表す入力の対応表

出力について、入力と同じように3つの値からなる列ベクトルを設定し、値が高い要素に対応するものをその出力と設定した。例えば出力ベクトル $\vec{o}$ が

$$\vec{o} = \begin{bmatrix} -1.2 \\ 0.8 \\ 2.4 \end{bmatrix}$$

となった場合、1番目の値はグー、2番目の値はチョキ、3番目の値はパーを表す。それぞれの値を見ると、-1.2, 0.8, 2.4となっており、3番目のパーに対応する値が一番大きくなっている。この場合の出力はパーであると見なす。なお、例えば1番目と2番目が同じ値で一番大きな値となった場合は、グーかチョキかのどちらかだということを示していることとする。

これらのことから、入力を図1に従って、グーを表す列ベクトルを入力すれば、出力の列ベクトルは3番目の値が一番大きくなればよく、図1に従ってチョキを表す列ベクトルを入力すれば、出力の列ベクトルは1番目の値が大きくなればよく、図1に従ってパーを表す列ベクトルを入力すれば、出力の列ベクトルは2番目の値が大きくなればよい。

このようなシステムを実現する際に、山田さんは、⁽²⁾ニューラルネットワークというアルゴリズムを知った。ニューラルネットワークは、ノードとエッジからなり、訓練データを用いて、エッジの重みを訓練することにより、決められた入力から望まれた出力を導出することを可能とするものである。

山田さんは、入力3つの値からなる列ベクトル、出力は3つの値からなる列ベクトルであることを踏まえ、中間層を設定することにより表現力が増し、正しい結果を得やすくなることを知り、中間層として5つの列ベクトルを設定し、それらが各層で繋がっている、図2のようなニューラルネットワークを構築した。

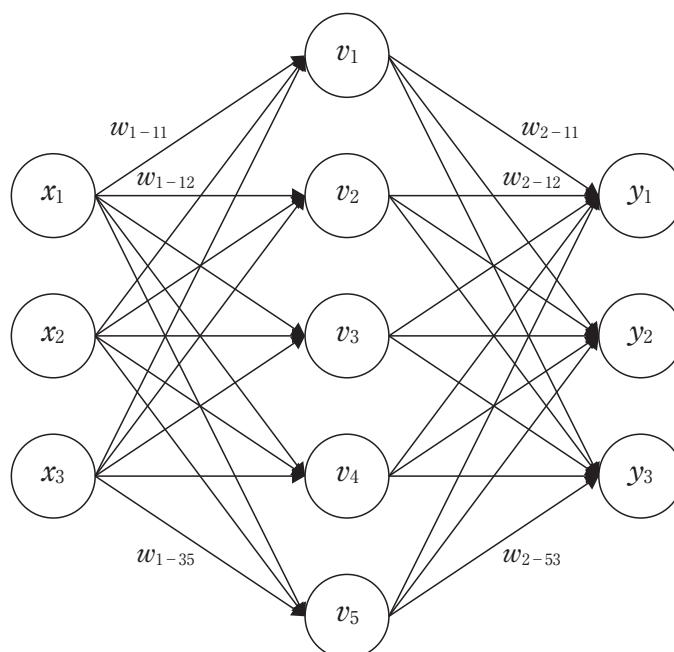


図2 山田さんが構成したニューラルネットワーク

このニューラルネットワークは入力層が x_1 , x_2 , x_3 、中間層が v_1 , v_2 , v_3 , v_4 , v_5 、出力層が y_1 , y_2 , y_3 からなる3層のニューラルネットワークである。 w_{h-ij} は、各層の要素をつなげるエッジの重みとなっており、例えば x_2 と v_3 をつなぐエッジの重みは w_{1-23} 、例えば v_4 と y_3 をつなぐエッジの重みは w_{2-43} と表現される。なお、このニューラルネットワークにおいては、活性化関数は考えなくて良い（活性化関数を知らない方は無視して構わない）。

入力層 x_1 , x_2 , x_3 に対して、グーを入力したい場合は、図1より、 $x_1=1$, $x_2=0$, $x_3=0$ のように入力を行う。その結果、出力層 y_1 , y_2 , y_3 の値については y_3 が一番大きい値になれば、本システムの要件を満たすこととなる。

ただし、ニューラルネットワークの初期状態では、入力層と中間層を結ぶ各エッジの重み w_{1-ij} 、中間層と出力層の各エッジの重み w_{2-jk} が決まっていない（もしくはランダムで

割り当てられている) ため, 正しい結果を得ることができない。

入力層と中間層を結ぶ各エッジの重み w_{1-ij} , 中間層と出力層の各エッジの重み w_{2-jk} をデータのパターンから調整することを訓練と呼ぶ。山田さんは図3のようなデータを用意した。

	データ 1	データ 2	データ 3
入力	1	0	0
	0	1	0
	0	0	1
出力	0	1	0
	0	0	1
	1	0	0

図3 山田さんが準備した訓練データ

図3のデータから確率的勾配降下法を使ってN回の学習をおこなった(本問題において確率的勾配降下法の詳細な仕組みを知る必要はない)。

N=5として学習をおこなった結果, 入力層と中間層を結ぶ各エッジの重み w_{1-ij} , 中間層と出力層の各エッジの重み w_{2-ij} の値が決定され, 次のような結果となった。

(3) $x_1=1, x_2=0, x_3=0$ のとき, $y_1=1.9205401, y_2=-3.9234746, y_3=0.7925714$ となり, グーという結果となった。
 $x_1=0, x_2=1, x_3=0$ のとき, $y_1=1.5776458, y_2=-3.3222334, y_3=1.0937274$ となり, グーという結果となった。
 $x_1=0, x_2=0, x_3=1$ のとき, $y_1=3.2520027, y_2=-5.2746296, y_3=0.9269992$ となり, グーという結果となった。

問1 下線部(1)について, グー, チョキ, パーをワンホットベクトル表現手法のようなものを用いて表現しなくてはならない理由について, 最も適切なものを, 下記の選択肢から選び, 番号をマークしなさい。☒ ア

選択肢

- ① グーとチョキとパーは指の数が0, 2, 5となっておりその大小関係を示すため
- ② アルゴリズムの中でグーとチョキとパーを一意の概念として数値で表現をするため
- ③ グーとチョキとパーは著作権があるため, これらをコピーすることを防ぐため
- ④ グーとチョキとパーを容易にソートすることによって, 計算量を減らすことが可能になるため

問2 図1の ～ について、適切な数値の0か1を選び、それぞれ番号をマークしなさい（同じ数値を選択しても構わない）。

問3 下線部（2）について、ニューラルネットワークのように、コンピュータにデータを訓練させ、データ内に潜むパターンを学習し、その学習結果を用いて、入力に対して学習結果をあてはめ、推定・予測する手法のことを何というか。下記の選択肢から選び、番号をマークしなさい。

選択肢

- ① 機械学習
- ② パターンマッチング
- ③ ファジー理論
- ④ ルールベース

問4 下線部（3）で導出された結果について、この場合の判断として最も適切ものを下記の選択肢から選び、番号をマークしなさい。

選択肢

- ① 出力結果は期待通りであったため、訓練はこれで終了として良い
- ② $N=5$ のまま、もう一度訓練を行う
- ③ $N=3$ として訓練を行う
- ④ $N=100$ として訓練を行う