Sample Problems and Answer Key

USA Junior AI Olympiad (USA-J-AIO)

Instructions

This exam contains 40 questions. You have 60 minutes.
Every question is multiple choice. Exactly one answer is correct.

Problem 1

A pet-photo app records five features in this order: pointed ears, collar, striped fur, visible tail, and indoors. It uses 1 for yes and 0 for no. An indoor cat has pointed ears and a collar, no stripes, and its tail is out of view. Which vector represents the photo?

A. $[1, 1, 0, 1, 0]$

B. $[1, 0, 1, 0, 1]$

C. $[1, 1, 0, 0, 1]$

D. $[0, 1, 0, 0, 1]$

E. $[1, 1, 1, 0, 1]$

Problem 2

A residual block computes $\mathbf{x}+F(\mathbf{x})$. The designer wants the block to pass its input through unchanged when the learned branch has no useful correction. What can $F(\mathbf{x})$ output to make that happen?

A. $\mathbf{x}$

B. $2\mathbf{x}$

C. $-\mathbf{x}$

D. The same nonzero constant for every input

E. The zero vector

Problem 3

Three projects train AI systems. (1) Pictures are paired with correct recycling-bin names. (2) Objects are grouped by similarity using unlabeled examples. (3) A game agent learns by receiving points for successful actions. Which matching is correct?

A. Unsupervised, supervised, reinforcement

B. Reinforcement, unsupervised, supervised

C. Supervised, reinforcement, unsupervised

D. Supervised, supervised, supervised

E. Supervised, unsupervised, reinforcement

Problem 4

A transit app predicts the number of minutes until a bus arrives and whether it will be crowded. Which statement is correct?

A. Arrival time is a regression target. Crowding is a classification target.

B. Both are regression targets because both concern buses.

C. Both are classification targets because both are predictions.

D. Arrival time is a classification target. Crowding is a regression target.

E. Both are regression targets if the app stores crowded as 1 and not crowded as 0.

Problem 5

A game character uses this artificial neuron:

```
def action_signal(energy, danger):
    number = 2 * energy - danger - 1
    return max(0, number)
```

Its energy is 2. Its danger reading increases from 5 to 6. What happens to the output?

A. It drops from -2 to -3.

B. It drops from 2 to 1.

C. It rises from 0 to 1.

D. It drops from 1 to 0.

E. It stays at 0.

Problem 6

During diffusion training, a program makes two noisy versions of the same clean image by sampling different noise each time. The network learns to predict the noise added in each example. Why can the program supply these training targets automatically?

A. Both versions use the clean image's object label as their noise target.

B. The program knows which noise it added to each version.

C. Both versions must contain the same noise because they share a clean image.

D. The network's first guess becomes the correct target.

E. The program assigns every pixel a target of zero.

Problem 7

A music app records how many songs a listener played in five genres. One listener has the vector $[2, 0, 4, 2, 6]$. Which other listener divides their plays among the five genres in exactly the same proportions?

A. $[3, 1, 5, 3, 7]$

B. $[1, 0, 2, 1, 3]$

C. $[4, 0, 2, 2, 6]$

D. $[2, 0, 4, 2, 5]$

E. $[2, 0, 4, 3, 5]$

Problem 8

A school club wants voice recordings to train a reading assistant. It wants students to understand how recordings will be used and be able to decline while still joining the same club activities. Which plan meets both goals?

A. Explain how recordings will be used, then require a recording from everyone who joins.

B. Explain the training use before recording, ask each student to opt in, and offer the same activities to students who decline.

C. Record everyone during the first activity, then ask permission before using the recordings for training.

D. Let students decline recording, but describe the purpose only as future school projects.

E. Ask students to opt in and explain the training use, but offer the main activities only to students who record.

Problem 9

A recycling robot uses this decision tree. If an item is metal, send it to recycling. Otherwise, send it to landfill if it is dirty and to recycling if it is clean.
The robot sees a dirty metal can and a dirty plastic tray. Both are then cleaned and scanned again. Which decision changes?

A. Both items change from landfill to recycling.

B. Only the can changes from landfill to recycling.

C. Only the tray changes from recycling to landfill.

D. Both decisions stay the same because neither material changed.

E. Only the tray changes from landfill to recycling.

Problem 10

A trained model should use the same learned weight each time it predicts.

```
def predict(x, weights):
    score = x * weights[0]
    weights[0] += 1
    return score

weights = [2]
print(predict(3, weights))
print(predict(3, weights))
```

What happens?

A. It prints 6 twice because the input is the same.

B. It prints 9 twice because the weight increases before either score is calculated.

C. It prints 9, then 12 because the new weight is used immediately in each call.

D. It prints 6, then 9 because the first prediction changes the weight used by the second.

E. It prints 6 twice because a function cannot change a list it receives.

Problem 11

Predictor P misses its delivery-time estimates by 0, 0, and 6 minutes. Predictor Q misses by 2, 2, and 2 minutes. The company changes its evaluation from the sum of absolute errors to the sum of squared errors. How does the comparison change?

A. P wins under both measures.

B. Q wins under both measures.

C. P wins under absolute error and Q wins under squared error.

D. They tie under absolute error and Q wins under squared error.

E. They tie under absolute error and P wins under squared error.

Problem 12

A trained model assigns risk scores to packages.

Package	P	Q	R	S
Risk score	0.9	0.7	0.6	0.4
Actually faulty?	Yes	No	Yes	No

```
def inspect(risk_score, threshold):
    return risk_score >= threshold
```

The team lowers the threshold from 0.8 to 0.5 and reuses the same scores. What changes?

A. One more faulty package is inspected, with no extra false alarms.

B. One more faulty package and one more good package are inspected. The scores stay the same.

C. Two more faulty packages are inspected because the model learns from the lower threshold.

D. One more good package is inspected, and the number of faulty packages caught stays the same.

E. All four packages are inspected because the threshold is lower.

Problem 13

A camera stores the rows of an image as lists inside `image`. A programmer wants to transpose it as shown.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

```
result = []
for c in range(3):
    new_row = []
    for r in range(2):
        new_row.append(________)
    result.append(new_row)
```

What belongs in the blank?

A. `image[r][c]`

B. `image[c][r]`

C. `image[0][c]`

D. `image[r][0]`

E. `image[1 - r][c]`

Problem 14

A recommender stores reader preferences in the rows of U and book features in the columns of V. Columns 1, 2, and 3 represent Books 1, 2, and 3.

$$U = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \qquad V = \begin{bmatrix} 2 & 1 & 2 \\ 0 & 2 & 0 \end{bmatrix}.$$

The product UV gives the scores. For example, Reader 1's score for Book 1 is $1 \cdot 2 + 1 \cdot 0 = 2$. The app updates U while keeping V fixed. Which statement is guaranteed?

A. Both readers will give Book 2 the same score.

B. Each reader will give Books 1 and 3 equal scores.

C. Book 1's score will stay the same for both readers.

D. Book 2 will become the highest-scoring book for both readers.

E. The two readers will give the same scores to all three books as each other.

Problem 15

A rover starts at the center of a flat map. The movement vector $[1, -1]$ means one unit right and one unit down. The rover can scale either of its two movement commands by any number, then add the movements. Which pair lets it reach any point on the map?

A. $[1, 1]$ and $[1, -1]$

B. $[1, 2]$ and $[2, 4]$

C. $[1, 0]$ and $[2, 0]$

D. $[0, 1]$ and $[0, -1]$

E. $[2, 2]$ and $[-1, -1]$

Problem 16

A classifier receives redness r and roundness s, in that order. Each row of its weight matrix gives the weights for one score.

$$W = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}.$$

For example, the first row gives $2r + s$. The code uses the rows of W as `weights`.

```
def scores(features, weights):
    output = []
    for row in weights:
        output.append(row[0] * features[0]
                      + row[1] * features[1])
    return output
```

New software sends the features in the order $[s, r]$. Which matrix keeps both scores the same, in their original order?

A. $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$

B. $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$

C. $\begin{bmatrix} 2 & 2 \\ 3 & 3 \end{bmatrix}$

D. $\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$

E. $\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$

Problem 17

A team tries 100 settings for a handwriting model. It chooses the setting with the highest validation accuracy, 90%. The validation examples were never used to calculate training updates. What should the team do for its final evaluation?

A. Report 90% as an untouched test result.

B. Adjust the model repeatedly until its test accuracy reaches 95%.

C. Rename the validation set “test set” and evaluate again.

D. Train on the validation examples, then report accuracy on those same examples.

E. Freeze its choice and evaluate on a separate, untouched test set.

Problem 18

A wildlife classifier takes a majority vote among its k nearest labeled photos. For one new photo, the five nearest labels, from closest to farthest, are fox, dog, dog, fox, fox. Which statement is correct?

A. The prediction is Fox for all three settings $k = 1, 3, 5$.

B. The predictions are Fox, Dog, and Dog for $k = 1, 3, 5$, respectively.

C. The predictions are Fox, Dog, and Fox, so $k = 5$ must work best on new photos.

D. The prediction is Dog at $k = 3$, so the closest photo must have the wrong label.

E. The predictions are Fox, Dog, and Fox. This example alone cannot tell us which setting is most accurate on new photos.

Problem 19

A team builds a classifier that must recognize new physical objects. It can collect either 1,000 video frames showing 10 objects or 500 frames showing 100 objects. Which plan best tests whether the model works on new objects?

A. Randomly split frames from the 10 objects so that each object appears in both sets.

B. Use the 100-object collection and reserve whole objects for testing.

C. Train on every frame in the 100-object collection and test on a random subset of those frames.

D. Reserve frames from the 10 objects, then use their test scores to choose when to stop training.

E. Compare the training accuracies and select whichever model scores higher.

Problem 20

A speech recognizer has these test results:

Speaker group	Recordings	Accuracy
Accent P	900	96%
Accent Q	100	60%

Accent Q was rare in training. The team can obtain transcripts for 100 more recordings. Which plan most directly addresses the gap and checks whether it improves?

A. Add recordings from Q, then evaluate on fresh examples from both groups separately.

B. Add recordings from P, then report only overall accuracy.

C. Add recordings from Q, then evaluate only on those same training recordings.

D. Add recordings from Q, then remove Q from the final test set.

E. Add recordings from P, then combine both groups into one reported score.

Problem 21

A robot updates its memory as it receives messages. Each triple gives how much old memory to keep, how much new information to write, and the candidate value.

```
memory = 4
gates = [(1, 0, 9), (0, 1, 2), (1, 0, 7)]

for keep, write, candidate in gates:
    memory = keep * memory + write * candidate
```

Which change leaves the final memory the same?

A. Change the first candidate from 9 to 99.

B. Change the second candidate from 2 to 20.

C. Change the second write amount from 1 to 0.

D. Change the last keep amount from 1 to 0.

E. Change the last write amount from 0 to 1.

Problem 22

A clustering program compares an object at $(0, 0)$ with centers A at $(0, 3)$ and B at $(2, 0)$.

```
def distance_squared(p, center):
    return ((p[0] - center[0]) ** 2
            + (p[1] - center[1]) ** 2)
```

The nearer center wins. The first coordinate is then expressed in a smaller unit, multiplying every first coordinate by 10. What happens?

A. The assignment changes from B to A.

B. The assignment changes from A to B.

C. A wins both times.

D. B wins both times.

E. Both comparisons are ties.

Problem 23

Two training runs use the same four good-part photos and four faulty-part photos. Each starts with weight $w = 5$. At that starting weight, each good-part photo contributes an update signal of $+1$, and each faulty-part photo contributes -3. The model adds the mean signal of its batch to w.
Run P's first batch has four good-part photos. Run Q's first batch has two good-part and two faulty-part photos. What are the weights after the first update, in the order (P, Q)?

A. $(6, 6)$

B. $(9, 1)$

C. $(4, 6)$

D. $(6, 4)$

E. $(5, 5)$

Problem 24

A model predicts a number using wx. A programmer tries to improve it:

```
w = -2
x = -2
target = 6

prediction = w * x
error = target - prediction
w = w + 0.1 * error
```

What happens when the prediction is recalculated using the new `w`?

A. It moves closer to 6 because `error` is positive.

B. It becomes exactly 6.

C. It moves farther from 6. Decreasing `w` to -3 would reach the target.

D. It stays unchanged because `x` is negative.

E. It becomes impossible to calculate.

Problem 25

A shop predicts tomorrow's sales at the end of day t. The list `sales` contains the entire study period, including future test days.

```
average = sum(sales) / len(sales)
features, targets = [], []

for t in range(1, len(sales) - 1):
    features.append([sales[t], average])
    targets.append(sales[t + 1])
```

Which change makes the average use only information available by day t?

A. Shuffle the entire list before calculating its average.

B. Round the entire list's average to a whole number.

C. Replace the average with tomorrow's sales.

D. Recalculate the entire list's average separately for each day.

E. Calculate each day's average using only sales from the beginning through day t.

Problem 26

A camera sees a dark box 2 pixels wide and 3 pixels tall, surrounded by a light wall. The six pixels with brightness 40 form the box. The filter examines the middle three rows and rightmost three columns, crossing the box's right edge.

$$\underbrace{\begin{bmatrix} 160 & 160 & 160 & 160 \\ 160 & 40 & 40 & 160 \\ 160 & 40 & 40 & 160 \\ 160 & 40 & 40 & 160 \\ 160 & 160 & 160 & 160 \end{bmatrix}}_{\text{Full image}} \qquad \underbrace{\begin{bmatrix} 40 & 40 & 160 \\ 40 & 40 & 160 \\ 40 & 40 & 160 \end{bmatrix}}_{\text{Edge patch}}$$

Brighter lighting adds 20 to every pixel. Which kernel gives a positive response that stays unchanged? Sum the products of corresponding pixel values and kernel weights. Use each kernel as shown, with zero bias.

A. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

B. $\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

C. $\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$

D. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

E. $\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$

Problem 27

Two strips of five light sensors record brightness readings $P = [0, 2, 0, 2, 0]$ and $Q = [0, 0, 2, 0, 0]$. A filter slides across adjacent pairs. It records the increase in brightness from left to right, using zero if brightness falls. For example, $[0, 2]$ gives 2 and $[2, 0]$ gives 0.
A classifier sees only one summary of the filter responses. Which statement is correct?

A. The maximum response distinguishes the strips because P has two bright spots.

B. The sum cannot distinguish the strips because their brightest readings are equal.

C. The maximum is the same for both, but the sum is larger for P because P has two rising edges.

D. Both summaries distinguish the strips.

E. Neither summary distinguishes the strips because the same filter is used.

Problem 28

Four sensors form a line, A − B − C − D. Their initial values are $A = 1$, $B = 2$, $C = 4$, and $D = 8$.

```
Repeat twice:
    Copy all current values into old.
    Set each sensor's new value to the sum of its neighbors' values in old.
    Replace all values simultaneously.
```

What is A's final value when D starts at 8? What would it be if only D's initial value changed to 16? Give the results in that order.

A. $(2, 2)$

B. $(5, 13)$

C. $(13, 21)$

D. $(5, 5)$

E. $(2, 10)$

Problem 29

A recycling camera counts red pixels x and blue pixels y. Each count is a whole number from 0 to 10. Two encoders store the following results:

$$P\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x+y\\x-y\end{bmatrix}, \qquad Q\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}1 & 1\\2 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x+y\\2x+2y\end{bmatrix}.$$

A second device receives the two stored numbers and knows which encoder was used. Which encoder lets it recover both original counts for every allowed pair?

A. Only P

B. Only Q

C. Both P and Q

D. Neither P nor Q

E. It depends on which count is larger.

Problem 30

An image-processing system repeatedly applies the same matrix to a feature vector:

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ x + 2y \end{bmatrix}.$$

It starts with $[2, 0]$. Each output becomes the input for the next application. Which statement remains true after every application?

A. The second coordinate stays zero.

B. The first coordinate is twice the second.

C. The first coordinate exceeds the second by exactly 2.

D. The two coordinates add up to 2.

E. Both coordinates double at every application.

Problem 31

An avatar editor uses these functions:

```
def stretch(point):
    x, y = point
    return [2 * x, y]

def turn(point):
    x, y = point
    return [-y, x]
```

For which starting points does `turn(stretch(point))` equal `stretch(turn(point))`?

A. Every point

B. Every point with $x = 0$

C. Only $[0, 0]$

D. Every point with $y = 0$

E. Every point with $x = y$

Problem 32

A sorting robot describes containers by width in centimetres, height in centimetres, whether they have a handle, and whether they have a lid, in that order. For the last two features, 1 means yes and 0 means no.

Feature vector	Label
$[10, 20, 0, 0]$	Type R
$[10, 20, 0, 1]$	Type S
$[30, 40, 1, 0]$	Type R
$[30, 40, 1, 1]$	Type S

Encoder P keeps width, height, and handle. Encoder Q keeps only lid. A classifier receives only the retained features and gives the same label for the same input. What are the highest possible accuracies on these four examples, in the order (P, Q)?

A. $(100\%, 50\%)$

B. $(100\%, 100\%)$

C. $(50\%, 50\%)$

D. $(50\%, 100\%)$

E. $(0\%, 100\%)$

Problem 33

A factory tests two detectors on 1,000 parts, including 10 faulty parts.

Detector	Faulty parts caught	Good parts wrongly flagged
P	0	0
Q	8	20

A missed fault costs 10 points. A false alarm costs 1 point. Which statement is correct?

A. P has both higher accuracy and lower cost.

B. Q has both higher accuracy and lower cost.

C. P has higher accuracy, but Q has lower cost.

D. Q has higher accuracy, but P has lower cost.

E. Both detectors have the same accuracy and cost.

Problem 34

An attention system combines three cards. If a card is blocked, its weight is removed and the remaining weights are rescaled to sum to 1.

Card	Value	Original weight
P	6	$1/2$
Q	0	$1/3$
R	6	$1/6$

The code uses the card order P, Q, R.

```
def read(values, weights, blocked):
    total = 0
    weight_sum = 0
    for i in range(len(values)):
        if i != blocked:
            total += weights[i] * values[i]
            weight_sum += weights[i]
    return total / weight_sum
```

A programmer argues that blocking Q cannot matter because Q's value is zero. What happens to the output when Q is blocked?

A. It stays at 4 because Q contributed zero to the original numerator.

B. It drops from 4 to 0 because Q's value is zero.

C. It rises from 4 to 6 because the remaining weight is shared entirely by cards with value 6.

D. It rises from 4 to 5 because one of the three cards was removed.

E. It stays at 6 because both remaining cards have value 6.

Problem 35

An autoencoder sees two pictures equally often, $P = [1, 0]$ and $Q = [0, 1]$. Here, 0 is black, 1 is white, and $1/2$ is gray. The encoder stores only the sum of the two pixels. The decoder receives only that sum and always gives the same output for the same sum. Which reconstruction minimizes average squared pixel error?

A. $[1, 0]$

B. $[0, 1]$

C. $[1, 1]$

D. $[1/2, 1/2]$

E. $[0, 0]$

Problem 36

A classifier must label $x = 0$ Red, $x = 1$ Blue, and $x = 2$ Red. Its network uses only linear transformations and biases, followed by one threshold decision. Which change allows correct classification of all three inputs?

A. Add more linear layers before the same type of threshold.

B. Train longer on copies of the same examples.

C. Use the nonlinear feature $|x - 1|$, predicting Blue when it is less than $1/2$.

D. Double x before the linear layers.

E. Add the same constant to every input before the linear layers.

Problem 37

A language model generates exactly two words. Its first word is cats with probability 60% or robots with probability 40%.

First word	P(sleep next)	P(dance next)
cats	60%	40%
robots	100%	0%

Greedy decoding chooses the most probable next word at each step. Which choice gives the greedy output followed by the most probable complete caption?

A. *cats sleep* and *robots sleep*

B. *cats sleep* and *cats sleep*

C. *robots sleep* and *robots sleep*

D. *cats dance* and *robots sleep*

E. *robots sleep* and *cats sleep*

Problem 38

An audit examines 1,000 submissions. Ten involved cheating. The AI detector flags 80% of the submissions involving cheating and 10% of the submissions without cheating. Which interpretation is correct?

A. 80% of flagged submissions involved cheating, so flags establish guilt.

B. 90% of flagged submissions involved cheating, so further review adds little.

C. Exactly 10% of flagged submissions involved cheating.

D. Fewer than 10% of flagged submissions involved cheating. A flag should prompt a review of the evidence before any penalty.

E. The proportion of correct flags cannot be calculated without inspecting the model's weights.

Problem 39

A Transformer uses the final representation at position 2 to predict the word at position 3. In this simplified model, each output is the weighted sum of the three representations from the preceding layer.

	Position 1	Position 2	Position 3
Layer 1 output at position 1	1/2	0	1/2
Layer 1 output at position 2	0	1/2	1/2
Layer 2 output at position 2	1/2	1/2	0

Which change prevents the original word at position 3 from reaching the prediction while keeping both of Layer 2's connections shown above?

A. Block Layer 1's connection from original position 3 to position 1 only.

B. Block Layer 1's connections from original position 3 to both positions 1 and 2.

C. Block Layer 1's connection from original position 3 to position 2 only.

D. Set the third weight in Layer 2's row to zero.

E. Make each row's weights sum to 1.

Problem 40

Three classifiers label the same ten photographs as cat or dog. Each gets exactly eight correct. Their predictions are combined by majority vote. What is the smallest possible number the combined system could get correct?

A. 6

B. 7

C. 8

D. 9

E. 10

Answer Key

Question	Answer	Question	Answer	Question	Answer	Question	Answer
1	C	11	D	21	A	31	C
2	E	12	B	22	A	32	D
3	E	13	A	23	D	33	C
4	A	14	B	24	C	34	C
5	E	15	A	25	E	35	D
6	B	16	D	26	E	36	C
7	B	17	E	27	C	37	A
8	B	18	E	28	D	38	D
9	E	19	B	29	A	39	B
10	D	20	A	30	C	40	B