Supplementary Material S1

Knowledge Test and Answer Key

From Users to Creators: Design and Preliminary Evaluation of an AI-Enabled Physical Systems Workshop

Instrument and translation

The following pages present an English translation of the Slovenian knowledge test administered immediately before and after the workshop. The translation retains the content of the questions and response options, the original figures, and the answer key. The same instrument and scoring criteria were used in all three implementations.

Quantitative scoring

The quantitative score is the sum of 16 multiple-choice items: one point for each correct response and zero points for an incorrect response, giving a maximum of 16 points. Question 11.3 is an additional open-ended task and was not included in the quantitative score or statistical analyses reported in the article.

Knowledge domain	Question identifiers	Maximum score
Electronics/programming	1, 2, 3, 4, 5, 6, 7	7
AI/ML/data	8, 9.1, 9.2, 9.3, 10.1, 10.2, 11.1, 11.2, 12	9

Item identifiers in the manuscript correspond directly to the question and subquestion numbers in this supplement (for example, Q9.1 refers to question 9.1). The answer key follows the complete test and includes an illustrative response to the open-ended task.

Knowledge Test: Use of Digital Technologies with Artificial Intelligence

Participant ID: _____ Date: _____

For multiple-choice questions, circle the single most appropriate option. For open-ended tasks, write your own answer or prompt as you would send it to an AI assistant.

1. What is shown in the image?



A) An expansion board with an H-bridge for controlling two motors.
B) A microcontroller development board with a camera and wireless connectivity.
C) An infrared sensor for detecting motion.
D) A device for capturing and storing images that can be connected to a computer.

2. What can you do with the Arduino IDE?

A) Design printed circuit boards and connections between components.
B) Train image-recognition models without writing code.
C) Design 3D models of enclosures and mechanical assemblies.
D) Write, verify, and upload programs to supported microcontrollers.

3. A push button has two states (pressed/not pressed). Which type of microcontroller pin is most suitable for detecting its state?

A) Digital output.
B) Analog output.
C) Digital input.
D) Analog input.

4. Which type of microcontroller pin is most suitable for turning an LED on and off through a program?

A) Digital output.
B) Digital input.
C) Analog input.
D) Analog output.

5. The anode of an LED is connected to pin D4 through a current-limiting resistor, and its cathode is connected to GND. Pin D4 is already configured as OUTPUT. Which command turns on the LED?

A) pinMode(4, HIGH);
B) digitalWrite(4, LOW);
C) digitalRead(4);
D) digitalWrite(4, HIGH);

6. Which C++ programming structure executes a block of code only when a specified condition is met (without repeating a loop)?

A) for
B) if
C) while
D) setup()

7. What happens to the electrical resistance of a light-dependent resistor (LDR) when less light falls on it?

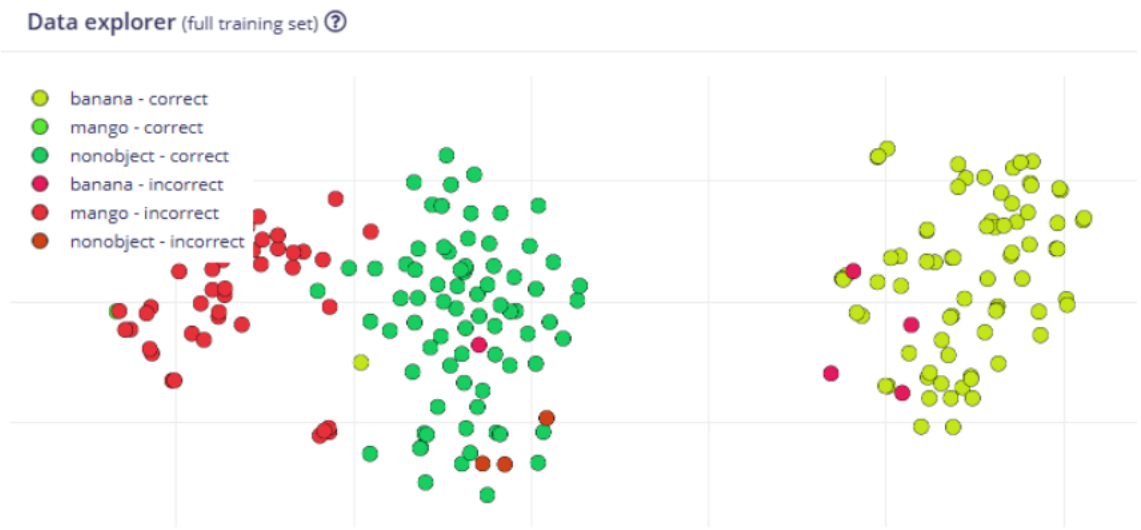
A) It increases.
B) It decreases.
C) It remains approximately unchanged.
D) It begins to oscillate periodically on its own.

8. You are collecting images to train a model to recognize different objects. Which data-collection approach is most appropriate to prevent the model from depending only on the recording conditions?

A) Always photograph each class from the same distance and against the same background.
B) Collect mainly a large number of nearly identical consecutive images for each class.
C) Include different viewpoints, backgrounds, and lighting conditions for each class.
D) Always photograph different classes against different backgrounds so the model can distinguish them more easily.

9. Understanding the visualization of training images

The Data Explorer visualization below shows images of bananas, mangoes, and other objects (nonobject) from the training dataset.



9.1 What does the label "incorrect" indicate for a point in the legend?

A) The image was captured under insufficient lighting.
B) The model's prediction does not match the image's correct label.
C) The model did not produce a prediction for the image.
D) The image was not included in the training dataset.

9.2 Which action would be the most appropriate first step toward reducing misclassifications?

A) Review the misclassified examples and their labels, and add diverse training images if necessary.
B) Remove all images that the model currently misclassifies.
C) Always photograph every object under the same lighting conditions and against the same background.
D) Add images only for the class that the model already recognizes most accurately.

9.3 What is the most justified conclusion about the group of banana images compared with the group of mango images?

A) All banana images were definitely taken from the same angle.
B) The model definitely used fewer banana images than mango images for training.
C) The model definitely makes no errors when recognizing bananas.
D) The banana images are more tightly clustered in the feature visualization; the graph does not reveal why.

10. Interpreting a confusion matrix

The matrix shows the model's results when classifying images into three classes. Rows represent the actual class, and columns represent the model's predictions.



10.1 What percentage of images that actually show bananas did the model classify as "nonobject"?

A) 6.7% **B)** 11.8% **C)** 88.2% **D)** 93.3%

10.2 What percentage of mango images did the model correctly recognize as mangoes?

A) 93.3% **B)** 6.7% **C)** 0% **D)** 100%

11. Comparing results after retraining the model

We added new training images to the model and retrained it. For the comparison, assume that the two matrices were obtained using comparable validation datasets. The second matrix also includes an "uncertain" prediction outcome.

Confusion matrix

ACTUAL LABEL \ PREDICTION LABEL	BANANA	MANGO	NONOBJECT	UNCERTAIN
BANANA	94.1%	0%	0%	5.9%
MANGO	0%	100%	0%	0%
NONOBJECT	10%	0%	80%	10%
F1 SCORE	0.91	1.00	0.89	

11.1 For which class did the percentage of correctly classified images increase the most?

A) Mango.
B) Nonobject.
C) Banana.
D) None of the classes.

11.2 For which class did the percentage of correctly classified images decrease?

A) Banana.
B) Mango.
C) None of the classes.
D) Nonobject.

11.3 Suggest one possible explanation for why the percentage of correctly classified mango images increased while that of the nonobject class decreased. How would you test your explanation?

12. We train a model using 100 images of cars, 20 images of boats, and 400 images of bicycles. Which statement is most justified based on the distribution of the training data?

A) The model will necessarily make the most errors when recognizing boats.
B) The number of images in each class cannot affect training.
C) There is a risk that the model will recognize the underrepresented boat class less accurately; performance must be evaluated separately for each class.
D) The model will learn the class with the fewest examples best because that class is rarer.

Answer Key

Multiple-choice questions

Question	Answer	Question	Answer
1	**B**	9.1	**B**
2	**D**	9.2	**A**
3	**C**	9.3	**D**
4	**A**	10.1	**B**
5	**D**	10.2	**C**
6	**B**	11.1	**A**
7	**A**	11.2	**D**
8	**C**	12	**C**

Question 11.3 - open-ended response

Example of an acceptable answer: Adding or changing training images may have improved recognition of mangoes while leaving the nonobject class insufficiently represented or less diverse. To test this explanation, examine the new training images and class labels, inspect misclassified nonobject examples, and compare class-specific results on the same or comparable validation data.

Other plausible explanations are acceptable if they are consistent with the confusion matrices and include a reasonable method of verification.