Question 1

An artist has been using an AI tool to create digital art and would like to ensure that it has copyright protection in the United States.

Which of the following is most likely to enable the artist to receive copyright protection?

Options:

A- Ensure the tool was trained using publicly available content.

B- Obtain a representation from the AI provider on how the tool works.

C- Provide a log of the prompts the artist used to generate the images.

D- Update the images in a creative way to demonstrate that it is the artist's.

Answer:

D

Explanation:

For the artist to receive copyright protection, the most effective approach is to demonstrate that the final artwork includes sufficient creative input by the artist. By updating or altering the images in a way that reflects the artist's personal creativity, the artist can claim originality, which is a core requirement for copyright protection under U.S. law. The other options do not directly address the originality and creative input required for copyright. This is highlighted in the sections on copyright protection in the IAPP AIGP Body of Knowledge.

Question 2

Question Type: **MultipleChoice**

A company plans on procuring a tool from an AI provider for its employees to use for certain business purposes.

Which contractual provision would best protect the company's intellectual property in the tool, including training and testing data?

Options:

A- The provider will give privacy notice to individuals before using their personal data to train or test the tool.

B- The provider will defend and indemnify the company against infringement claims.

C- The provider will obtain and maintain insurance to cover potential claims.

D- The provider will warrant that the tool will work as intended.

Answer:

B

Explanation:

To protect the company's intellectual property, the most pertinent contractual provision is ensuring that the AI provider will defend and indemnify the company against infringement claims. This clause means the provider will take responsibility for any intellectual property disputes that arise, thereby safeguarding the company from potential legal and financial repercussions related to the use of the tool. Other options, while beneficial, do not directly address the protection of intellectual property. This concept is detailed in the contractual best practices section of the IAPP AIGP Body of Knowledge.

Question 3

Question Type: **MultipleChoice**

Retraining an LLM can be necessary for all of the following reasons EXCEPT?

Options:

A- To minimize degradation in prediction accuracy due to changes in data.

B- Adjust the model's hyper parameters specific use case.

C- Account for new interpretations of the same data.

D- To ensure interpretability of the model's predictions.

Answer:

D

Explanation:

Retraining an LLM (Large Language Model) is primarily done to improve or maintain its performance as data changes over time, to fine-tune it for specific use cases, and to incorporate new data interpretations to enhance accuracy and relevance. However, ensuring interpretability of the model's predictions is not typically a reason for retraining. Interpretability relates to how easily the outputs of the model can be understood and explained, which is generally addressed through different techniques or methods rather than through the retraining process itself. References to this can be found in the IAPP AIGP Body of Knowledge discussing model retraining and interpretability as separate concepts.

Question 4

Question Type: **MultipleChoice**

You are an engineer that developed an AI-based ad recommendation tool.

Which of the following should be monitored to evaluate the tool's effectiveness?

Options:

A- Output data, assess the delta between the prediction and actual ad clicks.

B- Algorithmic patterns, to show the model has a high degree of accuracy.

C- Input data, to ensure the ads are reaching the target audience.

D- GPU performance, to evaluate the tool's robustness.

Answer:

A

Explanation:

To evaluate the effectiveness of an AI-based ad recommendation tool, the most relevant metric is the output data, specifically assessing the delta between the prediction and actual ad clicks. This metric directly measures the tool's accuracy and effectiveness in making accurate recommendations that lead to user engagement. While monitoring algorithmic patterns and input data can provide insights into the model's behavior and targeting accuracy, and GPU performance can indicate the robustness and efficiency of the tool, the primary indicator of effectiveness for an ad recommendation tool is how well it predicts actual ad clicks.

Question 5

You are part of your organization's ML engineering team and notice that the accuracy of a model that was recently deployed into production is deteriorating.

What is the best first step address this?

Options:

A- Replace the model with a previous version.

B- Conduct champion/challenger testing.

C- Perform an audit of the model.

D- Run red-teaming exercises.

Answer:

B

Explanation:

When the accuracy of a model deteriorates, the best first step is to conduct champion/challenger testing. This involves deploying a new model (challenger) alongside the current model (champion) to compare their performance. This method helps identify if the new model can perform better under current conditions without immediately discarding the existing model. It provides a controlled environment to test improvements and understand the reasons behind the deterioration. This approach is preferable to directly replacing the model, performing audits, or running red-teaming exercises, which may be subsequent steps based on the findings from the champion/challenger testing.

Question 6

Question Type: MultipleChoice

All of the following are reasons to deploy a challenger AI model in addition a champion AI model EXCEPT to?

Options:

A- Provide a framework to consider alternatives to the champion model.

B- Automate real-time monitoring of the champion model.

C- Perform testing on the champion model.

D- Retrain the champion model.

Answer:

D

Explanation:

Deploying a challenger AI model alongside a champion model is a strategy used to compare the performance of different models in a real-world environment. This approach helps in providing a framework to consider alternatives to the champion model, automating real-time monitoring of the champion model, and performing testing on the champion model. However, retraining the champion model is not a reason to deploy a challenger model. Retraining is a separate process that involves updating the champion model with new data or techniques, which is not related to the use of a challenger model.

Question 7

Question Type: MultipleChoice

All of the following are included within the scope of post-deployment AI maintenance EXCEPT?

Options:

A- Ensuring that all model components are subject a control framework.

B- Dedicating experts to continually monitor the model output.

C- Evaluating the need for an audit under certain standards.

D- Defining thresholds to conduct new impact assessments.

Answer:

D

Explanation:

Post-deployment AI maintenance typically includes ensuring that all model components are subject to a control framework, dedicating experts to continually monitor the model output, and evaluating the need for audits under certain standards. However, defining thresholds to conduct new impact assessments is usually part of the initial deployment and ongoing governance processes rather than a maintenance activity. Maintenance focuses more on the operational aspects of the AI system rather than setting new thresholds for impact assessments.

Question 8

Question Type: **MultipleChoice**

What is the best method to proactively train an LLM so that there is mathematical proof that no specific piece of training data has more than a negligible effect on the model or its output?

Options:

A- Clustering.

B- Transfer learning.

C- Differential privacy.

D- Data compartmentalization.

Answer:

C

Explanation:

Differential privacy is a technique used to ensure that the inclusion or exclusion of a single data point does not significantly affect the outcome of any analysis, providing a way to mathematically prove that no specific piece of training data has more than a negligible effect on the model or its output. This is achieved by introducing randomness into the data or the algorithms processing the data. In the context of training large language models (LLMs), differential privacy helps in protecting individual data points while still enabling the model to learn effectively. By adding noise to the training process, differential privacy provides strong guarantees about the privacy of the training data.

Question 9

What is the primary purpose of conducting ethical red-teaming on an AI system?

Options:

A- To improve the model's accuracy.

B- To simulate model risk scenarios.

C- To identify security vulnerabilities.

D- To ensure compliance with applicable law.

Answer:

B

Explanation:

The primary purpose of conducting ethical red-teaming on an AI system is to simulate model risk scenarios. Ethical red-teaming involves rigorously testing the AI system to identify potential weaknesses, biases, and vulnerabilities by simulating real-world attack or failure scenarios. This helps in proactively addressing issues that could compromise the system's reliability, fairness, and security. Reference: AIGP Body of Knowledge on AI Risk Management and Ethical AI Practices.

Question 10

Question Type: MultipleChoice

What is the technique to remove the effects of improperly used data from an ML system?

Options:

A- Data cleansing.

B- Model inversion.

C- Data de-duplication.

D- Model disgorgement.

Answer:

D

Explanation:

Model disgorgement is the technique used to remove the effects of improperly used data from an ML system. This process involves retraining or adjusting the model to eliminate any biases or inaccuracies introduced by the inappropriate data. It ensures that the model's outputs are not influenced by data that was not meant to be used or was used incorrectly. Reference: AIGP Body of Knowledge on Data Management and Model Integrity.