Practice With CompTIA CY0-001 Mock Test

Shared by Cervantes on 17-06-2026

For More Free Questions and Preparation Resources

Check the Links on Last Page

Question 1

Question Type: MultipleChoice

Which option best explains the reason a cybersecurity analyst prefers a machine learning (ML) model over a statistical model for attack classification?

Options:

A- The ability to learn complex problems and adapt to new information
B- A simplified development pipeline and deployment process
C- Improved performance with a small data set and high durability
D- Large community support and availability of global experts

Answer:

A

Explanation:

Basic Concept: Cybersecurity threats evolve continuously, with new attack variants emerging regularly. The choice between traditional statistical models and ML models for attack classification depends on which better handles the complexity and dynamism of the threat landscape. CompTIA SecAI+ covers ML model advantages for cybersecurity under basic AI concepts.

Why A is Correct: ML models can learn arbitrarily complex, non-linear relationships from training data and adapt to new patterns when retrained with updated data. For attack classification, this means ML can recognize sophisticated, multi-feature attack patterns that exceed the capabilities of simple statistical models and can be updated to detect new attack variants as the threat landscape evolves. This adaptability to complex and changing problems is the primary reason analysts prefer ML over static statistical approaches.

Why B is Wrong: ML model development pipelines are generally more complex than statistical models, requiring data preparation, feature engineering, model selection, training, validation, and deployment steps. Simplicity of development is not a characteristic advantage of ML over statistical models.

Why C is Wrong: ML models typically require large amounts of training data to perform well. Statistical models often perform better than ML with small datasets. Performance with small datasets is actually an advantage of statistical models over ML, not ML over statistical.

Why D is Wrong: Community support and expert availability are ecosystem considerations rather than technical reasons to prefer ML for cybersecurity classification tasks. These factors might

influence tool selection but do not explain the fundamental technical preference for ML's superior handling of complex attack patterns.

Question 2

Question Type: MultipleChoice

Which of the following attacks is most enabled by AI-generated content?

Options:

A- Model poisoning
B- Phishing
C- Ransomware
D- Remote code execution

Answer:

B

Explanation:

Basic Concept: AI-generated content including personalized text, synthetic voice, and deepfake video has dramatically enhanced the effectiveness and scalability of social engineering attacks. Understanding how AI amplifies specific attack types is key to CompTIA SecAI+ basic AI concepts in the cybersecurity context.

Why B is Correct: Phishing attacks are most dramatically enabled by AI-generated content. AI can generate highly personalized, grammatically perfect phishing emails tailored to individual targets using publicly available information. It can create convincing deepfake audio and video for voice phishing (vishing) and video phishing, replicate executive communication styles for business email compromise, and generate phishing campaigns at massive scale. The quality and personalization that previously required skilled human social engineers can now be automated with AI.

Why A is Wrong: Model poisoning is a specific attack against AI systems that corrupts training data to manipulate model behavior. While sophisticated, it is a targeted AI security attack rather than a broad cybercrime enabled by AI-generated content at scale.

Why C is Wrong: Ransomware is malware that encrypts victim data and demands payment for decryption keys. While AI can assist in ransomware development, ransomware deployment relies on code execution and network propagation techniques more than AI-generated content.

Why D is Wrong: Remote code execution involves exploiting vulnerabilities to run arbitrary code on a target system. It relies on technical vulnerability exploitation rather than AI-generated content. AI might assist in finding vulnerabilities, but RCE is not primarily enabled by content generation.

Question 3

Question Type: MultipleChoice

A line of business wants to onboard an application that uses a custom AI model for employee assessments. The Chief Information Officer (CIO) agrees to allow the engagement to proceed but first wants a threat model.

Which of the following is the best to use for an AI threat model?

Options:

A- Responsible AI
B- Adversarial Threat Landscape for AI Systems (ATLAS)
C- Organization for Economic Co-operation and Development (OECD)
D- International Organization for Standardization (ISO)

Answer:

B

Explanation:

Basic Concept: Threat modeling for AI systems requires a framework specifically designed to address AI-specific attack techniques, tactics, and procedures. General cybersecurity or governance frameworks do not capture the unique adversarial attack surface of AI and ML systems. CompTIA SecAI+ Exam Objectives identify MITRE ATLAS as the primary AI threat modeling resource.

Why B is Correct: MITRE ATLAS (Adversarial Threat Landscape for AI Systems) is specifically designed as an AI and ML threat modeling framework. It catalogs real-world adversarial tactics, techniques, and procedures targeting AI systems, enabling security architects to identify and assess threats unique to ML models such as data poisoning, model extraction, and evasion attacks. It is the industry standard for AI-specific threat modeling.

Why A is Wrong: Responsible AI is a set of ethical principles and governance guidelines for developing and deploying AI systems fairly and safely. It addresses ethics and fairness, not

technical adversarial threat modeling.

Why C is Wrong: The OECD provides non-binding policy recommendations and principles for AI governance at an international level. It does not provide technical threat modeling taxonomies or AI-specific attack catalogs.

Why D is Wrong: ISO standards such as ISO 42001 establish management system requirements for AI governance. They are compliance and management frameworks, not threat modeling tools for identifying adversarial AI attack vectors.

Question 4

Question Type: MultipleChoice

A customer-facing, AI-powered chatbot has been jailbroken through prompt injections. As a result, the AI model is offering a 99% discount on the purchase of a new vehicle.

Which of the following should be implemented to enhance the model's robustness against such attacks?

Options:

A- Bias filtering
B- System prompt
C- Log monitoring
D- Guardrails

Answer:

D

Explanation:

Basic Concept: Jailbreaking through prompt injection exploits the LLM's tendency to follow instructions embedded in user input, overriding its intended behavior. The model was manipulated to offer unauthorized discounts, demonstrating that its operational boundaries were not properly enforced. CompTIA SecAI+ Study Guide identifies guardrails as the primary defense against jailbreaking attacks.

Why D is Correct: Guardrails are robust, layered controls that enforce behavioral boundaries on LLM inputs and outputs. They can detect and block jailbreaking attempts, enforce business logic constraints such as preventing unauthorized discounts, validate outputs against policy rules

before delivery, and prevent the model from operating outside its defined scope. Guardrails are specifically designed to make models more robust against prompt injection and jailbreaking.

Why A is Wrong: Bias filtering is designed to detect and remove biased, discriminatory, or offensive content from model outputs. It addresses content fairness issues but does not prevent jailbreaking attacks that manipulate the model into performing unauthorized actions.

Why B is Wrong: A system prompt sets the model's base instructions and persona, but the jailbreak attack already demonstrates that the current prompt can be overridden. Guardrails provide enforcement at a layer that is more resistant to prompt manipulation than the system prompt alone.

Why C is Wrong: Log monitoring detects jailbreaking attempts after they have already succeeded. It is a detective control that enables incident response but does not prevent the model from offering unauthorized discounts in the first place.

Question 5

Question Type: MultipleChoice

A company develops an AI model to diagnose patients. Hospitals access the model through an integrated application programming interface (API). The security team performs a denial-of-service (DoS) attack via brute force on the model.

Which of the following controls would have prevented this issue?

Options:

A- Tokenization
B- Model guardrails
C- Rate limiting
D- Prompt firewall

Answer:

C

Explanation:

Basic Concept: API-based AI systems are susceptible to DoS attacks where excessive requests overwhelm the system's ability to respond to legitimate users. Rate limiting is the standard control for preventing both intentional and unintentional API abuse. CompTIA SecAI+ Study Guide

covers rate limiting as a key availability control for AI APIs.

Why C is Correct: Rate limiting restricts the number of requests a client can make to an API within a defined time window. In this scenario, a brute-force DoS attack works by sending a massive volume of requests to exhaust the model's resources. Rate limiting would have automatically throttled or blocked the excessive request volume, preventing the attack from succeeding and preserving service availability for legitimate hospital users.

Why A is Wrong: Tokenization replaces sensitive data values with non-sensitive placeholders. It is a data security control for protecting sensitive information such as patient identifiers, not a control for managing API request volumes or preventing DoS attacks.

Why B is Wrong: Model guardrails filter and constrain model inputs and outputs for safety and policy compliance. They inspect content quality, not request volume, and cannot prevent a volume-based DoS attack.

Why D is Wrong: A prompt firewall inspects the content of prompts for malicious patterns or policy violations. Like guardrails, it analyzes content rather than controlling request frequency and cannot prevent resource exhaustion from a high-volume brute-force attack.

Question 6

Question Type: MultipleChoice

A security operations center (SOC) has a very high volume of logs and alerts. The manager proposes the implementation of a machine learning (ML) system to help with triage.

Which of the following tasks is most suitable?

Options:

A- Applying filters on specific alerts
B- Automatically patching vulnerable systems
C- Identifying and classifying alerts
D- Summarizing the content of alerts

Answer:

C

Explanation:

Basic Concept: ML models excel at classification tasks, learning to assign incoming data points to predefined categories based on patterns in training data. In a SOC context, alert classification is the highest-value triage function ML can perform. CompTIA SecAI+ Exam Objectives address AI-assisted security operations under Domain 3.

Why C is Correct: ML-based alert classification automatically analyzes characteristics of each alert and assigns it to a severity category such as critical, high, medium, or low, or to a threat type such as malware or intrusion attempt. This dramatically reduces analyst workload and speeds triage by prioritizing which alerts demand immediate human attention, directly solving the high-volume problem.

Why A is Wrong: Applying filters on specific alerts is a rule-based operation achievable without ML using simple log management tools. It requires no learning capability and does not adapt to new or evolving threats.

Why B is Wrong: Automatically patching systems is a remediation action requiring validated, controlled processes. Having an ML system autonomously patch production systems without human oversight poses unacceptable operational and security risk.

Why D is Wrong: Summarizing alert content is a useful generative AI function but does not provide prioritization value for triage. Classification tells analysts what to act on first; summarization only rephrases existing information.

Question 7

Question Type: MultipleChoice

Which of the following controls is the best way to mitigate a denial-of-service (DoS) attack?

Options:

A- Model guardrails
B- Rate limiting
C- End-to-end encryption
D- Access controls

Answer:

B

Explanation:

Basic Concept: DoS attacks overwhelm AI systems by sending excessive requests that exhaust computational resources, memory, or bandwidth, preventing legitimate users from being served. The primary defense against volume-based attacks is throttling the rate at which requests can be processed. CompTIA SecAI+ Exam Objectives identify rate limiting as the key DoS mitigation control for AI systems.

Why B is Correct: Rate limiting directly addresses the root mechanism of DoS attacks by restricting the number of requests any single client or IP address can submit within a defined time window. By enforcing request quotas, rate limiting prevents attackers from generating the request volume necessary to overwhelm the system while preserving capacity for legitimate users. It is the most direct and effective preventive control against DoS attacks on AI APIs and services.

Why A is Wrong: Model guardrails inspect and filter the content of prompts and responses for policy compliance and safety. They operate at the semantic content level, not at the request volume level, and cannot prevent resource exhaustion from high-volume request flooding.

Why C is Wrong: End-to-end encryption protects the confidentiality and integrity of data in transit. Encrypted DoS traffic is just as damaging as unencrypted traffic; encryption does not limit request rates or prevent resource exhaustion.

Why D is Wrong: Access controls restrict who can interact with the system, which can reduce the potential attacker pool. However, authenticated users and compromised accounts can still launch DoS attacks, and access controls alone cannot prevent high-volume attacks from authorized sources.

Question 8

Question Type: MultipleChoice

A team of engineers builds an application using a large language model (LLM). The application is built on Linux and is hosted on a virtual server. Users must create an account in order to access and use the platform.

Which of the following should the team do to protect the account credentials?

Options:

A- Patch the model with the latest data set.
B- Update the Linux and virtual servers.
C- Implement hashing and encryption.
D- Deploy an authenticated application programming interface (API).

Answer:

C

Explanation:

Basic Concept: User account credentials stored in a database must be protected against unauthorized disclosure. The security of credentials at rest requires cryptographic controls that prevent even database administrators or attackers with database access from reading plaintext passwords. CompTIA SecAI+ Study Guide covers credential security controls as part of AI application security.

Why C is Correct: Implementing hashing and encryption for credential protection is the industry-standard approach. Passwords should be hashed using strong, slow algorithms such as bcrypt, Argon2, or scrypt with unique salts, making them computationally infeasible to reverse even if the database is compromised. Additional sensitive credential data can be encrypted. Together, hashing and encryption ensure that account credentials remain protected even if the underlying storage is accessed by unauthorized parties.

Why A is Wrong: Patching the model with new datasets updates the AI model's training data and knowledge. It does not address the security of user account credentials stored in the application's authentication database.

Why B is Wrong: Updating Linux and virtual server software patches system vulnerabilities and is important for overall security hygiene. However, it does not implement specific protections for the account credentials themselves stored in the application database.

Why D is Wrong: Deploying an authenticated API requires users to authenticate to use the API, improving access control. While this complements credential security, it does not protect the storage of credentials at rest and does not replace hashing and encryption of the credential values themselves.

Question 9

Question Type: MultipleChoice

A user interface engineer adds new graphics to the latest release of an AI-integrated application. During the update, the engineer accidentally causes the model to retrain on unverified dat

a. After the update, the model begins to return many errors.

Which of the following is the best way to mitigate future errors?

Options:

A- Web application firewall
B- Role-based access control
C- Model development life cycle
D- Generative adversarial network

Answer:

C

Explanation:

Basic Concept: When a non-ML engineer can accidentally trigger model retraining during a UI update, this indicates a lack of proper lifecycle management and change controls around the AI model. Uncontrolled retraining on unverified data is a critical vulnerability in the development and deployment process. CompTIA SecAI+ Study Guide identifies the Model Development Life Cycle as the framework for preventing such issues.

Why C is Correct: Implementing a Model Development Life Cycle (MDLC) establishes formal, controlled processes for every stage of model development and updates including data validation requirements before training, change management gates, testing and validation stages, and separation of duties between UI development and model training activities. An MDLC would have prevented the accidental retraining by requiring explicit, controlled authorization before any model training occurs.

Why A is Wrong: A WAF filters HTTP traffic at the application boundary. It does not govern internal development processes or control when and how model retraining occurs within the AI development pipeline.

Why B is Wrong: Role-based access control can restrict who has permission to trigger model retraining, which would help prevent this specific incident. However, it is one component of a broader MDLC governance framework and does not address data validation, testing stages, or the complete change management process.

Why D is Wrong: A GAN is a model architecture for generating synthetic data. It is a training technique unrelated to lifecycle governance or preventing accidental retraining from unverified data during unrelated application updates.

Question 10

Question Type: MultipleChoice

An attacker successfully completes a denial-of-service (DoS) attack through the context window of an AI system. Thousands of characters are obfuscated and hidden behind an emoji.

Which of the following techniques best mitigates this type of attack?

Options:

A- Fraud detection
B- Large language model (LLM)-as-a-judge
C- Pattern recognition
D- Prompt filter

Answer:

D

Explanation:

Basic Concept: Context window DoS attacks flood an LLM's context with obfuscated content to exhaust processing resources or manipulate model behavior. Attackers may hide large amounts of text behind Unicode characters like emojis. CompTIA SecAI+ Study Guide identifies prompt filtering as the primary defense against input-based attacks on LLMs.

Why D is Correct: A prompt filter inspects incoming inputs before they reach the LLM, detecting and blocking malicious content including obfuscated text hidden behind Unicode characters or emojis. By analyzing input structure, character counts, hidden content, and encoding anomalies, prompt filters can identify and reject attacks that attempt to abuse the context window, preventing resource exhaustion.

Why A is Wrong: Fraud detection systems are designed to identify fraudulent transactions or activities in structured data contexts. They are not designed to inspect LLM prompt structures for obfuscated content attacks on context windows.

Why B is Wrong: LLM-as-a-judge uses a secondary LLM to evaluate the quality or safety of another model's outputs. It operates post-generation and cannot prevent a DoS attack that occurs during input processing before output is generated.

Why C is Wrong: Pattern recognition can identify known attack patterns but requires the attack to match pre-learned patterns. Novel obfuscation techniques using Unicode or emoji hiding may evade pattern-based detection without dedicated prompt filtering logic.

Question 11

Question Type: MultipleChoice

A security architect performs threat modeling of an AI system. The architect needs to determine which attacks can be performed against the system.

Which of the following actions should the architect take next?

Options:

A- Leverage a large language model (LLM) to map likely attack paths based on the code base.
B- Quantify the risk of known vulnerabilities identified in the AI system.
C- Identify trust boundaries and perform threat modeling with Open Worldwide Application Security Project (OWASP) Top 10.
D- Analyze MITRE Adversarial Threat Landscape for AI Systems (ATLAS) for tactics, techniques, and procedures (TTPs).

Answer:

D

Explanation:

Basic Concept: AI-specific threat modeling requires consulting resources that catalogue adversarial attacks specifically developed for AI and ML systems. General cybersecurity frameworks may miss AI-unique attack vectors such as model inversion, data poisoning, and adversarial examples. CompTIA SecAI+ Study Guide identifies MITRE ATLAS as the authoritative source for AI system TTPs.

Why D is Correct: MITRE ATLAS provides a comprehensive, curated knowledge base of adversarial tactics, techniques, and procedures specifically targeting AI and ML systems, derived from real-world attack case studies. Analyzing ATLAS enables the architect to enumerate realistic AI-specific attacks applicable to the system being threat-modeled, which directly answers the question of which attacks can be performed.

Why A is Wrong: Using an LLM to map attack paths introduces uncertainty and potential hallucination risk. LLMs may generate plausible-sounding but inaccurate attack paths and cannot guarantee comprehensive coverage of AI-specific attack techniques.

Why B is Wrong: Quantifying risk of known vulnerabilities is a risk assessment step that occurs after identifying which attacks are possible. The architect must first identify attack possibilities before quantifying their risk impact.

Why C is Wrong: OWASP Top 10 covers web application vulnerabilities and, in its LLM edition, certain LLM-specific risks. However, MITRE ATLAS provides a more comprehensive and structured catalog of AI and ML-specific adversarial TTPs for systematic threat modeling.

Question 12

Question Type: MultipleChoice

An organization is developing and implementing AI features into a customer service application.

Which option best practices should the organization put in place before releasing the application for customer trials?

Options:

A- Data masking and sanitization
B- External compliance audits
C- Approved AI vendor lists
D- Third-party risk management

Answer:

A

Explanation:

Basic Concept: Before deploying AI applications that handle customer data in trials, protecting sensitive information through data masking and sanitization is essential. CompTIA SecAI+ Study Guide emphasizes pre-deployment data security controls as a critical step in the AI development lifecycle.

Why A is Correct: Data masking replaces sensitive real customer data with realistic but fictitious equivalents, while sanitization removes harmful or unwanted data elements. Before customer trials, these techniques prevent exposure of real PII or sensitive information, ensure the trial environment cannot leak production data, and protect the organization from privacy regulation violations. This is the most immediately actionable pre-trial security control.

Why B is Wrong: External compliance audits are formal processes typically conducted post-deployment or at planned intervals to verify regulatory compliance. They are not pre-trial security implementations and cannot prevent data exposure in a trial environment.

Why C is Wrong: Approved AI vendor lists are governance artifacts that manage vendor selection

risk at the procurement stage. They do not directly protect customer data within an application being prepared for trials.

Why D is Wrong: Third-party risk management addresses risks from external vendors and partners at a strategic level. While important for overall governance, it does not constitute a direct data security control for a pre-trial release.