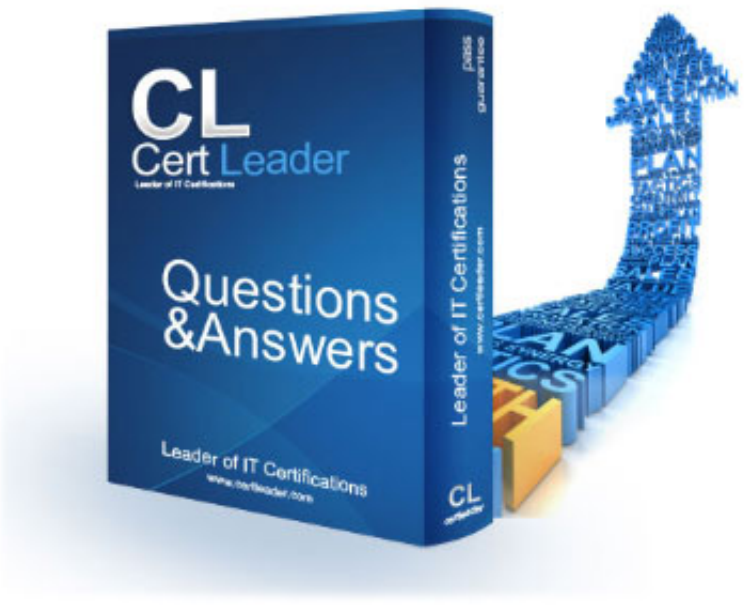
AIP-C01 Dumps

AWS Certified Generative AI Developer - Professional

https://www.certleader.com/AIP-C01-dumps.html



NEW QUESTION 1
A company is developing a generative AI (GenAI) application that uses Amazon Bedrock foundation models. The application has several custom tool integrations. The application has experienced unexpected token consumption surges despite consistent user traffic.
The company needs a solution that uses Amazon Bedrock model invocation logging to monitor InputTokenCount and OutputTokenCount metrics. The solution must detect unusual patterns in tool usage and identify which specific tool integrations cause abnormal token consumption. The solution must also automatically adjust thresholds as traffic patterns change.
Which solution will meet these requirements?

A. Use Amazon CloudWatch Logs to capture model invocation log
B. Create CloudWatch dashboards for token metric
C. Configure static CloudWatch alarms with fixed thresholds for each tool integration.
D. Store model invocation logs in Amazon S3. Use AWS Glue and Amazon Athena to analyze token usage trends.
E. Use Amazon CloudWatch Logs to capture model invocation log
F. Create CloudWatch metric filters to extract tool-specific invocation pattern
G. Apply CloudWatch anomaly detection alarms that automatically adjust baselines for each tool??s token metrics.
H. Store model invocation logs in an Amazon S3 bucke
I. Use AWS Lambda to process logs in real tim
J. Manually update CloudWatch alarm thresholds based on trends identified by the Lambda function.

Answer: C

NEW QUESTION 2
An ecommerce company is developing a generative AI application that uses Amazon Bedrock with Anthropic Claude to recommend products to customers. Customers report that some recommended products are not available for sale on the website or are not relevant to the customer. Customers also report that the solution takes a long time to generate some recommendations.
The company investigates the issues and finds that most interactions between customers and the product recommendation solution are unique. The company confirms that the solution recommends products that are not in the company??s product catalog. The company must resolve these issues.
Which solution will meet this requirement?

A. Increase grounding within Amazon Bedrock Guardrail
B. Enable Automated Reasoningcheck
C. Set up provisioned throughput.
D. Use prompt engineering to restrict the model responses to relevant product
E. Use streaming techniques such as the InvokeModelWithResponseStream action to reduce perceived latency for the customers.
F. Create an Amazon Bedrock knowledge bas
G. Implement Retrieval Augmented Generation RA
H. Set the PerformanceConfigLatency parameter to optimized.
I. Store product catalog data in Amazon OpenSearch Servic
J. Validate the model??s product recommendations against the product catalo
K. Use Amazon DynamoDB to implement response caching.

Answer: C

NEW QUESTION 3
A financial services company is deploying a generative AI (GenAI) application that uses Amazon Bedrock to assist customer service representatives to provide personalized investment advice to customers. The company must implement a comprehensive governance solution that follows responsible AI practices and meets regulatory requirements.
The solution must detect and prevent hallucinations in recommendations. The solution must have safety controls for customer interactions. The solution must also monitor model behavior drift in real time and maintain audit trails of all prompt-response pairs for regulatory review. The company must deploy the solution within 60 days. The solution must integrate with the company's existing compliance dashboard and respond to customers within 200 ms.
Which solution will meet these requirements with the LEAST operational overhead?

A. Configure Amazon Bedrock guardrails to apply custom content filters and toxicity detectio
B. Use Amazon Bedrock Model Evaluation to detect hallucination
C. Store prompt- response pairs in Amazon DynamoDB to capture audit trails and set a TT
D. Integrate Amazon CloudWatch custom metrics with the existing compliance dashboard.
E. Deploy Amazon Bedrock and use AWS PrivateLink to access the application securel
F. Use AWS Lambda functions to implement custom prompt validatio
G. Store prompt-response pairs in an Amazon S3 bucket and configure S3 Lifecycle policie
H. Create custom Amazon CloudWatch dashboards to monitor model performance metrics.
I. Use Amazon Bedrock Agents and Amazon Bedrock Knowledge Bases to ground response
J. Use Amazon Bedrock Guardrails to enforce content safet
K. Use Amazon OpenSearch Service to store and index prompt-response pair
L. Integrate OpenSearch Service with Amazon QuickSight to create compliance reports and to detect model behavior drift.
M. Use Amazon SageMaker Model Monitor to detect model behavior drif
N. Use AWS WAF to filter conten
O. Store customer interactions in an encrypted Amazon RDS databas
P. Use Amazon API Gateway to create custom HTTP APIs to integrate with the compliance dashboard.

Answer: A

NEW QUESTION 4
A company is building a legal research AI assistant that uses Amazon Bedrock with an Anthropic Claude foundation model (FM). The AI assistant must retrieve highly relevant case law documents to augment the FM??s responses. The AI assistant must identify semantic relationships between legal concepts, specific legal terminology, and citations. The AI assistant must perform quickly and return precise results.
Which solution will meet these requirements?

A. Configure an Amazon Bedrock knowledge base to use a default vector search configuratio

B. Use Amazon Bedrock to expand queries to improve retrieval for legal documents based on specific terminology and citations.
C. Use Amazon OpenSearch Service to deploy a hybrid search architecture that combines vector search with keyword searc
D. Apply an Amazon Bedrock reranker model to optimize result relevance.
E. Enable the Amazon Kendra query suggestion feature for end user
F. Use Amazon Bedrock to perform post-processing of search results to identify semantic similarity in the documents and to produce precise results.
G. Use Amazon OpenSearch Service with vector search and Amazon Bedrock Titan Embeddings to index and search legal document
H. Use custom AWS Lambda functions to merge results with keyword-based filters that are stored in an Amazon RDS database.

Answer: B

NEW QUESTION 5
A company uses an AI assistant application to summarize the company??s website content and provide information to customers. The company plans to use Amazon Bedrock to give the application access to a foundation model (FM).
The company needs to deploy the AI assistant application to a development environment and a production environment. The solution must integrate the environments with the FM. The company wants to test the effectiveness of various FMs in each environment. The solution must provide product owners with the ability to easily switch between FMs for testing purposes in each environment.
Which solution will meet these requirements?

A. Create one AWS CDK applicatio
B. Create multiple pipelines in AWS CodePipelin
C. Configure each pipeline to have its own settings for each F
D. Configure the application to invoke the Amazon Bedrock FMs by using the aws_bedrock.ProvisionedModel.fromProvisionedModelArn() method.
E. Create a separate AWS CDK application for each environmen
F. Configure the applications to invoke the Amazon Bedrock FMs by using the aws_bedrock.FoundationModel.fromFoundationModelId() metho
G. Create a separate pipeline in AWS CodePipeline for each environment.
H. Create one AWS CDK applicatio
I. Configure the application to invoke the Amazon Bedrock FMs by using the aws_bedrock.FoundationModel.fromFoundationModelId() metho
J. Create a pipeline in AWS CodePipeline that has a deployment stage for each environment that uses AWS CodeBuild deploy actions.
K. Create one AWS CDK application for the production environmen
L. Configure the application to invoke the Amazon Bedrock FMs by using the aws_bedrock.ProvisionedModel.fromProvisionedModelArn() metho
M. Create a pipeline in AWS CodePipelin
N. Configure the pipeline to deploy to the production environment by using an AWS CodeBuild deploy actio
O. For the development environment, manually recreate the resources by referring to the production application code.

Answer: C

NEW QUESTION 6
A healthcare company is developing a document management system that stores medical research papers in an Amazon S3 bucket. The company needs a comprehensive metadata framework to improve search precision for a GenAI application. The metadata must include document timestamps, author information, and research domain classifications.
The solution must maintain a consistent metadata structure across all uploaded documents and allow foundation models (FMs) to understand document context without accessing full content.
Which solution will meet these requirements?

A. Store document timestamps in Amazon S3 system metadat
B. Use S3 object tags for domain classificatio
C. Implement custom user-defined metadata to store author information.
D. Set up S3 Object Lock with legal holds to track document timestamp
E. Use S3 object tags for author informatio
F. Implement S3 access points for domain classification.
G. Use S3 Inventory reports to track timestamp
H. Create S3 access points for domain classificatio
I. Store author information in S3 Storage Lens dashboards.
J. Use custom user-defined metadata to store author informatio
K. Use S3 Object Lock retention periods for timestamp
L. Use S3 Event Notifications for domain classification.

Answer: A

NEW QUESTION 7
A company is using Amazon Bedrock to design an application to help researchers apply for grants. The application is based on an Amazon Nova Pro foundation model (FM). The application contains four required inputs and must provide responses in a consistent text format. The company wants to receive a notification in Amazon Bedrock if a response contains bullying language. However, the company does not want to block all flagged responses.
The company creates an Amazon Bedrock flow that takes an input prompt and sends it to the Amazon Nova Pro FM. The Amazon Nova Pro FM provides a response.
Which additional steps must the company take to meet these requirements? (Select TWO.)

A. Use Amazon Bedrock Prompt Management to specify the required inputs as variable
B. Select an Amazon Nova Pro F
C. Specify the output format for the respons
D. Add the prompt to the prompts node of the flow.
E. Create an Amazon Bedrock guardrail that applies the hate content filte
F. Set the filter response to bloc
G. Add the guardrail to the prompts node of the flow.
H. Create an Amazon Bedrock prompt route
I. Specify an Amazon Nova Pro F
J. Add the required inputs as variables to the input node of the flo
K. Add the prompt router to the prompts nod
L. Add the output format to the output node.
M. Create an Amazon Bedrock guardrail that applies the insults content filte

N. Set the filter response to detec
O. Add the guardrail to the prompts node of the flow.
P. Create an Amazon Bedrock application inference profile that specifies an Amazon Nova Pro F
Q. Specify the output format for the response in the descriptio
R. Include a tag for each of the input variable
S. Add the profile to the prompts node of the flow.

Answer: AD

NEW QUESTION 8
A healthcare company is using Amazon Bedrock to develop a real-time patient care AI assistant to respond to queries for separate departments that handle clinical inquiries, insurance verification, appointment scheduling, and insurance claims. The company wants to use a multi-agent architecture.
The company must ensure that the AI assistant is scalable and can onboard new features for patients. The AI assistant must be able to handle thousands of parallel patient interactions. The company must ensure that patients receive appropriate domain-specific responses to queries.
Which solution will meet these requirements?

A. Isolate data for each agent by using separate knowledge base
B. Use IAM filtering to control access to each knowledge bas
C. Deploy a supervisor agent to perform natural language intent classification on patient inquirie
D. Configure the supervisor agent to route queries to specialized collaborator agents to respond to department-specific querie
E. Configure each specialized collaborator agent to use Retrieval Augmented Generation(RAG) with the agent's department-specific knowledge base.
F. Create a separate supervisor agent for each departmen
G. Configure individual collaborator agents to perform natural language intent classification for each specialty domain within each departmen
H. Integrate each collaborator agent with department-specific knowledge bases onl
I. Implement manual handoff processes between the supervisor agents.
J. Isolate data for each department in separate knowledge base
K. Use IAM filtering to control access to each knowledge bas
L. Deploy a single general-purpose agen
M. Configure multiple action groups within the general-purpose agent to perform specific department function
N. Implement rule-based routing logic in the general-purpose agent instructions.
O. Implement multiple independent supervisor agents that run in parallel to respond to patient inquiries for each departmen
P. Configure multiple collaborator agents for each supervisor agen
Q. Integrate all agents with the same knowledge bas
R. Use external routing logic to merge responses from multiple supervisor agents.

Answer: A

NEW QUESTION 9
A healthcare company uses Amazon Bedrock to deploy an application that generates summaries of clinical documents. The application experiences inconsistent response quality with occasional factual hallucinations. Monthly costs exceed the company??s projections by 40%. A GenAI developer must implement a near real-time monitoring solution to detect hallucinations, identify abnormal token consumption, and provide early warnings of cost anomalies. The solution must require minimal custom development work and maintenance overhead.
Which solution will meet these requirements?

A. Configure Amazon CloudWatch alarms to monitor InputTokenCount and OutputTokenCount metrics to detect anomalie
B. Store model invocation logs in an Amazon S3 bucke
C. Use AWS Glue and Amazon Athena to identify potential hallucinations.
D. Run Amazon Bedrock evaluation jobs that use LLM-based judgments to detect hallucination
E. Configure Amazon CloudWatch to track token usag
F. Create an AWS Lambda function to process CloudWatch metric
G. Configure the Lambda function to send usage pattern notifications.
H. Configure Amazon Bedrock to store model invocation logs in an Amazon S3 bucke
I. Enable text output loggin
J. Configure Amazon Bedrock guardrails to run contextual grounding checks to detect hallucination
K. Create Amazon CloudWatch anomaly detection alarms for token usage metrics.
L. Use AWS CloudTrail to log all Amazon Bedrock API call
M. Create a custom dashboard in Amazon QuickSight to visualize token usage pattern
N. Use Amazon SageMaker Model Monitor to detect quality drift in generated summaries.

Answer: C

NEW QUESTION 10
Company configures a landing zone in AWS Control Tower. The company handles sensitive data that must remain within the European Union. The company must use only the eu-central-1 Region. The company uses Service Control Policies (SCPs) to enforce data residency policies. GenAI developers at the company are assigned IAM roles that have full permissions for Amazon Bedrock.
The company must ensure that GenAI developers can use the Amazon Nova Pro model through Amazon Bedrock only by using cross-Region inference (CRI) and only in eu- central-1. The company enables model access for the GenAI developer IAM roles in Amazon Bedrock. However, when a GenAI developer attempts to invoke the model through the Amazon Bedrock Chat/Text playground, the GenAI developer receives the following error:
User arn:aws:sts:123456789012:assumed-role/AssumedDevRole/DevUserName Action: bedrock:InvokeModelWithResponseStream
On resource(s): arn:aws:bedrock:eu-west-3::foundation-model/amazon.nova-pro-v1:0 Context: a service control policy explicitly denies the action
The company needs a solution to resolve the error. The solution must retain the company's existing governance controls and must provide precise access control. The solution must comply with the company's existing data residency policies.
Which combination of solutions will meet these requirements? (Select TWO.)

A. Add an AdministratorAccess policy to the GenAI developer IAM role
B. Extend the existing SCPs to enable CRI for the eu.amazon.nova-pro-v1:0 inference profile
C. Enable Amazon Bedrock model access for Amazon Nova Pro in the eu-west-3 Region
D. Validate that the GenAI developer IAM roles have permissions to invoke Amazon Nova Pro through the eu.amazon.nova-pro-v1:0 inference profile on all European Union AWS Regions that can serve the model
E. Extend the existing SCP to enable CRI for the eu-* inference profile

Answer: BE

NEW QUESTION 10
A company is building a generative AI (GenAI) application that uses Amazon Bedrock APIs to process complex customer inquiries. During peak usage periods, the application experiences intermittent API timeouts that cause issues such as broken response chunks and delayed data delivery. The application struggles to ensure that prompts remain within token limits when handling complex customer inquiries of varying lengths. Users have reported truncated inputs and incomplete responses. The company has also observed foundation model (FM) invocation failures.
The company needs a retry strategy that automatically handles transient service errors and prevents overwhelming Amazon Bedrock during peak usage periods. The strategy must also adapt to changing service availability and support response streaming and token- aware request handling.
Which solution will meet these requirements?

A. Implement a standard retry strategy that uses a 1-second fixed delay between attempts and a 3-retry maximum for all error
B. Handle streaming response timeouts by restarting stream
C. Cap token usage for each session.
D. Implement an adaptive retry strategy that uses exponential backoff with jitter and a circuit breaker pattern that temporarily disables retries when error rates exceed a predefined threshol
E. Implement a streaming response handler that monitors for chunk delivery timeout
F. Configure the handler to buffer successfully received chunks and intelligently resume streaming from the last received chunk when connections are re-established.
G. Use the AWS SDK to configure a retry strategy in standard mod
H. Wrap Amazon Bedrock API calls in try-catch blocks that handle timeout exception
I. Return cached completions for failed streaming request
J. Enforce a global token limit for all user
K. Add jitter-based retry logic and lightweight token trimming for each reques
L. Resume broken streams by requesting only missing chunks from the point of failur
M. Maintain a small in- memory buffer of the most recent chunks.
N. Set Amazon Bedrock client request timeouts to 30 second
O. Implement client-side load sheddin
P. Buffer partial results and stop new requests when application performance degrade
Q. Set static token usage caps for all request
R. Configure exponential backoff retries, dynamic chunk sizing, and context-aware token limits.

Answer: B

NEW QUESTION 14
A company upgraded its Amazon Bedrock–powered foundation model (FM) that supports a multilingual customer service assistant. After the upgrade, the assistant exhibited inconsistent behavior across languages. The assistant began generating different responses in some languages when presented with identical questions. The company needs a solution to detect and address similar problems for future updates. The evaluation must be completed within 45 minutes for all supported languages. The evaluation must process at least 15,000 test conversations in parallel. The evaluation process must be fully automated and integrated into the CI/CD pipeline. The solution must block deployment if quality thresholds are not met.
Which solution will meet these requirements?

A. Create a distributed traffic simulation framework that sends translation-heavy workloads to the assistant in multiple languages simultaneousl
B. Use Amazon CloudWatch metrics to monitor latency, concurrency, and throughpu
C. Run simulations before production releases to identify infrastructure bottlenecks.
D. Deploy the assistant in multiple AWS Regions with Amazon Route 53 latency-based routing and AWS Global Accelerator to improve global performanc
E. Store multilingual conversation logs in Amazon S3. Perform weekly post-deployment audits to review consistency.
F. Create a pre-processing pipeline that normalizes all incoming messages into a consistent format before sending the messages to the assistan
G. Apply rule-based checks to flag potential hallucinations in the output
H. Focus evaluation on normalized text to simplify testing across languages.
I. Set up standardized multilingual test conversations with identical meanin
J. Run the test conversations in parallel by using Amazon Bedrock model evaluation job
K. Apply similarity and hallucination threshold
L. Integrate the process into the CI/CD pipeline to block releases that fail.

Answer: D

NEW QUESTION 19
A retail company has a generative AI (GenAI) product recommendation application that uses Amazon Bedrock. The application suggests products to customers based on browsing history and demographics. The company needs to implement fairness evaluation across multiple demographic groups to detect and measure bias in recommendations between two prompt approaches. The company wants to collect and monitor fairness metrics in real time. The company must receive an alert if the fairness metrics show a discrepancy of more than 15% between demographic groups. The company must receive weekly reports that compare the performance of the two prompt approaches.
Which solution will meet these requirements with the LEAST custom development effort?

A. Configure an Amazon CloudWatch dashboard to display default metrics from Amazon Bedrock API call
B. Create custom metrics based on model output
C. Set up Amazon EventBridge rules to invoke AWS Lambda functions that perform post-processing analysis on model responses and publish custom fairness metrics.
D. Create the two prompt variants in Amazon Bedrock Prompt Managemen
E. Use Amazon Bedrock Flows to deploy the prompt variants with defined traffic allocatio
F. Configure Amazon Bedrock guardrails to monitor demographic fairnes
G. Set up Amazon CloudWatch alarms on the GuardrailContentSource dimension by using InvocationsIntervened metrics to detect recommendation discrepancy threshold violations.
H. Set up Amazon SageMaker Clarify to analyze model output
I. Publish fairness metrics to Amazon CloudWatc
J. Create CloudWatch composite alarms that combine SageMaker Clarify bias metrics with Amazon Bedrock latency metrics.
K. Create an Amazon Bedrock model evaluation job to compare fairness between the two prompt variant
L. Enable model invocation logging in Amazon CloudWatc
M. Set up CloudWatch alarms for InvocationsIntervened metrics with a dimension for each demographic group.

Answer: B

NEW QUESTION 22
An ecommerce company operates a global product recommendation system that needs to switch between multiple foundation models (FM) in Amazon Bedrock based on regulations,
cost optimization, and performance requirements. The company must apply custom controls based on proprietary business logic, including dynamic cost thresholds, AWS Region-specific compliance rules, and real-time A/B testing across multiple FMs.
The system must be able to switch between FMs without deploying new code. The system must route user requests based on complex rules including user tier, transaction value, regulatory zone, and real-time cost metrics that change hourly and require immediate propagation across thousands of concurrent requests.
Which solution will meet these requirements?

A. Deploy an AWS Lambda function that uses environment variables to store routing rules and Amazon Bedrock FM ID
B. Use the Lambda console to update the environment variables when business requirements chang
C. Configure an Amazon API Gateway REST API to read request parameters to make routing decisions.
D. Deploy Amazon API Gateway REST API request transformation templates to implement routing logic based on request attribute
E. Store Amazon Bedrock FM endpoints as REST API stage variable
F. Update the variables when the system switches between models.
G. Configure an AWS Lambda function to fetch routing configurations from the AWS AppConfig Agent for each user reques
H. Run business logic in the Lambda function to select the appropriate FM for each reques
I. Expose the FM through a single Amazon API Gateway REST API endpoint.
J. Use AWS Lambda authorizers for an Amazon API Gateway REST API to evaluate routing rules that are stored in AWS AppConfi
K. Return authorization contexts based on business logi
L. Route requests to model-specific Lambda functions for each Amazon Bedrock FM.

Answer: C

NEW QUESTION 26
A wildlife conservation agency operates zoos globally. The agency uses various sensors, trackers, and audiovisual recorders to monitor animal behavior. The agency wants to launch a generative AI (GenAI) assistant that can ingest multimodal data to study animal behavior.
The GenAI assistant must support natural language queries, avoid speculative behavioral interpretations, and maintain audit logs for ethical research audits.
Which solution will meet these requirements?

A. Ingest raw videos into Amazon Rekognition to detect animal postures and expression
B. Use Amazon Data Firehose to stream sensor and GPS data into Amazon S3. Prompt an Amazon Bedrock FM using basic templates stored in AWS Systems Manager Parameter Stor
C. Use IAM for access contro
D. Use AWS CloudTrail for audit logging.
E. Use Amazon SageMaker Processing and Amazon Transcribe to pre-process multimodal dat
F. Ingest curated summaries into an Amazon Bedrock Knowledge Base
G. Apply Amazon Bedrock guardrails to restrict speculative output
H. Use AWS AppConfig to manage prompt template
I. Use AWS CloudTrail to log research activity for audits.
J. Use Amazon OpenSearch Serverless to index behavioral logs and telemetr
K. Use Amazon Comprehend to extract entitie
L. Use Amazon Bedrock to answer questions over indexed dat
M. Use IAM for access control and CloudTrail for audit logging.
N. Configure Amazon O Business to federate data across Amazon S3, Amazon Kinesis, and Amazon SageMaker Feature Stor
O. Use EventBridge for ingestion orchestratio
P. Use custom AWS Lambda functions to filter LLM outputs for ethical compliance.

Answer: B

NEW QUESTION 30
A company is developing a generative AI (GenAI)-powered customer support application that uses Amazon Bedrock foundation models (FMs). The application must maintain conversational context across multiple interactions with the same user. The application must run clarification workflows to handle ambiguous user queries. The company must store encrypted records of each user conversation to use for personalization. The application must be able to handle thousands of concurrent users while responding to each user quickly.
Which solution will meet these requirements?

A. Use an AWS Step Functions Express workflow to orchestrate conversation flo
B. Invoke AWS Lambda functions to run clarification logi
C. Store conversation history in Amazon RDS and use session IDs as the primary key.
D. Use an AWS Step Functions Standard workflow to orchestrate clarification workflow
E. Include Wait for a Callback patterns to manage the workflow
F. Store conversation history in Amazon DynamoD
G. Purchase on-demand capacity and configure server-side encryption.
H. Deploy the application by using an Amazon API Gateway REST API to route user requests to an AWS Lambda function to update and retrieve conversation contex
I. Store conversation history in Amazon S3 and configure server-side encryptio
J. Save each interaction as a separate JSON file.
K. Use AWS Lambda functions to call Amazon Bedrock inference API
L. Use Amazon SQS queues to orchestrate clarification step
M. Store conversation history in an Amazon ElastiCache (Redis OSS) cluste
N. Configure encryption at rest.

Answer: B

NEW QUESTION 34
A company is using Amazon Bedrock to develop a customer support AI assistant. The AI assistant must respond to customer questions about their accounts. The

AI assistant must not expose personal information in responses. The company must comply with data residency policies by ensuring that all processing occurs within the same AWS Region where each customer is located.
The company wants to evaluate how effective the AI assistant is at preventing the exposure of personal information before the company makes the AI assistant available to customers.
Which solution will meet these requirements?

A. Configure a cross-Region Amazon Bedrock guardrail to apply sensitive information filter
B. Set the guardrail to detect mode during development and testin
C. Switch to block mode for production deployment.
D. Configure an Amazon Bedrock guardrail to apply sensitive information filter
E. Set theguardrail to mask mode during development and testin
F. Switch to block mode for production deploymen
G. Deploy a copy of the guardrail to each Region where the company operates.
H. Configure an Amazon Bedrock guardrail to apply content and topic filter
I. Set the guardrail to detect mode during development, testing, and productio
J. Disable invocation logging for the Amazon Bedrock model.
K. Configure a cross-Region Amazon Bedrock guardrail to apply a set of content and word filter
L. Set the guardrail to detect mode during development and testin
M. Switch to mask mode for production deployment.

Answer: B

NEW QUESTION 38
A pharmaceutical company is developing a Retrieval Augmented Generation application that uses an Amazon Bedrock knowledge base. The knowledge base uses Amazon OpenSearch Service as a data source for more than 25 million scientific papers. Users report that the application produces inconsistent answers that cite irrelevant sections of papers when queries span methodology, results, and discussion sections of the papers.
The company needs to improve the knowledge base to preserve semantic context across related paragraphs on the scale of the entire corpus of data.
Which solution will meet these requirements?

A. Configure the knowledge base to use fixed-size chunkin
B. Set a 300-token maximum chunk size and a 10% overlap between chunk
C. Use an appropriate Amazon Bedrock embedding model.
D. Configure the knowledge base to use hierarchical chunkin
E. Use parent chunks that contain 1,000 tokens and child chunks that contain 200 token
F. Set a 50-token overlap between chunks.
G. Configure the knowledge base to use semantic chunkin
H. Use a buffer size of 1 and a breakpoint percentile threshold of 85% to determine chunk boundaries based on content meaning.
I. Configure the knowledge base not to use chunkin
J. Manually split each document into separate files before ingestio
K. Apply post-processing reranking during retrieval.

Answer: B

NEW QUESTION 43
A financial services company is developing a generative AI (GenAI) application that serves both premium customers and standard customers. The application uses AWS Lambda functions behind an Amazon API Gateway REST API to process requests. The company needs to dynamically switch between AI models based on which customer tier each user belongs to. The company also wants to perform A/B testing for new features without redeploying code. The company needs to validate model parameters like temperature and maximum token limits before applying changes.
Which solution will meet these requirements with the LEAST operational overhead?

A. Create AWS Systems Manager Parameter Store parameters for each configuratio
B. Use Lambda functions to poll for parameter update
C. Use Amazon EventBridge events to trigger redeployments when configurations change.
D. Store model configurations in Amazon DynamoDB table
E. Optimize access patterns to retrieve configurations according to customer tie
F. Configure Lambda functions to query DynamoDB at the beginning of each request to determine which model to use.
G. Use AWS AppConfig to manage model configuration
H. Use feature flags to perform A/B testin
I. Define JSON schema validation rules for model parameter
J. Configure Lambda functions to retrieve configurations by using the AWS AppConfig Agent.
K. Create an Amazon ElastiCache (Redis OSS) cluster to store model configuration
L. Set short TTL value
M. Run custom validation logic in Lambda function
N. Use Amazon CloudWatch metrics to monitor configuration usage.

Answer: C

NEW QUESTION 48
A company is building a video analysis platform on AWS. The platform will analyze a large video archive by using Amazon Rekognition and Amazon Bedrock. The platform must comply with predefined privacy standards. The platform must also use secure model I/O, control foundation model (FM) access patterns, and provide an audit of who accessed what and when.
Which solution will meet these requirements?

A. Configure VPC endpoints for Amazon Bedrock model API call
B. Implement Amazon Bedrock guardrails to filter harmful or unauthorized content in prompts and response
C. Use Amazon Bedrock trace events to track all agent and model invocations for auditing purpose
D. Export the traces to Amazon CloudWatch Logs as an audit record of model usag
E. Store all prompts and outputs in Amazon S3 with server-side encryption with AWS KMS keys (SSE-KMS).
F. Define access control by using IAM with attribute-based access control (ABAC) to map departments to specific permission
G. Configure VPC endpoints for Amazon Bedrock model API call

H. Use IAM condition keys to enforce specific GuardrailIdentifier and ModelId value
I. Configure AWS CloudTrail to capture management and data events for S3 objects and KMS key usage activitie
J. Enable S3 server access logging to record detailed file-level interactions with the video archive
K. Send all CloudTrail logs to AWS CloudTrail Lak
L. Set up Amazon CloudWatch alarms to detect and alert on unexpected activity from Amazon Bedrock, Amazon Rekognition, and AWS KMS.
M. Restrict access to services by using VPC endpoint policie
N. Use AWS Config to track resource changes and compliance with security rule
O. Use server-side encryption with AWS KMS keys (SSE-KMS) to encrypt data at res
P. Store the model??s I/O in separate Amazon S3 bucket
Q. Enable S3 server access logging to track file-level interactions.
R. Configure AWS CloudTrail Insights to analyze API call patterns across accounts and detect anomalous activity in Amazon Bedrock, Amazon Rekognition, Amazon S3, and AWS KM
S. Deploy Amazon Macie to scan and classify the video archiv
T. Use server-side encryption with AWS KMS keys (SSE-KMS) to encrypt all stored dat
. Configure CloudTrail to capture KMS API usage events for audit purpose
. Configure Amazon EventBridge rules to process CloudTrail Insights anomalies and Macie finding
. Use CloudWatch alarms to trigger automated notifications and security responses when potential security issues are detected.

Answer: B

NEW QUESTION 50
A company is planning to deploy multiple generative AI (GenAI) applications to five independent business units that operate in multiple countries in Europe and the Americas.
Each application uses Amazon Bedrock Retrieval Augmented Generation (RAG) patterns with business unit-specific knowledge bases that store terabytes of unstructured data.
The company must establish well-architected, standardized components for security controls, observability practices, and deployment patterns across all the GenAI applications. The components must be reusable, versioned, and governed consistently.
Which solution will meet these requirements?

A. Configure Amazon API Gateway REST API endpoints for the GenAI application
B. Deploy common security, observability, and RAG patterns based on the AWS Well- Architected Generative AI Lens in standardized AWS CloudFormation template
C. Use CloudFormation Guard after deployment to validate policy compliance in each business unit.
D. Create standardized AWS CloudFormation templates to implement security, observability, and RAG patterns based on the AWS Well-Architected Generative AI Len
E. Establish a centralized repository for version contro
F. Integrate a CI/CD pipeline with CloudFormation Guard to enforce consistent and repeatable deployments across business units.
G. Use AWS Service Catalog to define standardized portfolios and versioned products for each business uni
H. Use the portfolios to enforce security, observability, and RAG patterns based on the AWS Well-Architected Generative AI Len
I. Require business units to use the Service Catalog console to deploy resources.
J. Document security controls, observability requirements, and RAG patterns based on the AWS Well-Architected Generative AI Lens in a shared design documen
K. Use Amazon Macie to enforce deploymen
L. Delegate implementation responsibility to each business unit.

Answer: B

NEW QUESTION 54
A finance company is developing an AI assistant to help clients plan investments and manage their portfolios. The company identifies several high-risk conversation patterns such as requests for specific stock recommendations or guaranteed returns. High-risk conversation patterns could lead to regulatory violations if the company cannot implement appropriate controls.
The company must ensure that the AI assistant does not provide inappropriate financial advice, generate content about competitors, or make claims that are not factually grounded in the company's approved financial guidance. The company wants to use Amazon Bedrock Guardrails to implement a solution.
Which combination of steps will meet these requirements? (Select THREE)

A. Add the high-risk conversation patterns to a denied topics guardrail.
B. Configure a content filter guardrail to filter prompts that contain the high-risk conversation patterns.
C. Configure a content filter guardrail to filter prompts that contain competitor names.
D. Add the names of competitors as custom word filter
E. Set the input and output actions to block.
F. Set a low grounding score threshold.
G. Set a high grounding score threshold.

Answer: ADF

NEW QUESTION 57
A company wants to select a new FM for its AI assistant. A GenAI developer needs to generate evaluation reports to help a data scientist assess the quality and safety of various foundation models FMs. The data scientist provides the GenAI developer with sample prompts for evaluation. The GenAI developer wants to use Amazon Bedrock to automate report generation and evaluation.
Which solution will meet this requirement?

A. Combine the sample prompts into a single JSON documen
B. Create an Amazon Bedrock knowledge base with the documen
C. Write a prompt that asks the FM to generate a response to each sample promp
D. Use the RetrieveAndGenerate API to generate a report for each model.
E. Combine the sample prompts into a single JSONL documen
F. Store the document in an Amazon S3 bucke
G. Create an Amazon Bedrock evaluation job that uses a judge mode
H. Specify the S3 location as input and a different S3 location as outpu
I. Run an evaluation job for each FM and select the FM as the generator.
J. Combine the sample prompts into a single JSONL documen

K. Store the document in an Amazon S3 bucke
L. Create an Amazon Bedrock evaluation job that uses a judge mode
M. Specify the S3 location as input and Amazon QuickSight as outpu
N. Run an evaluation job for each FM and select the FM as the evaluator.
O. Combine the sample prompts into a single JSON documen
P. Create an Amazon Bedrock knowledge base from the documen
Q. Create an Amazon Bedrock evaluation job that uses the retrieval and response generation evaluation typ
R. Specify an Amazon S3 bucket as the outpu
S. Run an evaluation job for each FM.

Answer: B

NEW QUESTION 59
A healthcare company is developing an application to process medical queries. The application must answer complex queries with high accuracy by reducing semantic dilution. The application must refer to domain-specific terminology in medical documents to reduce ambiguity in medical terminology. The application must be able to respond to 1,000 queries each minute with response times less than 2 seconds.
Which solution will meet these requirements with the LEAST operational overhead?

A. Use Amazon API Gateway to route incoming queries to an Amazon Bedrock agen
B. Configure the agent to use an Anthropic Claude model to decompose queries and an Amazon Titan model to expand querie
C. Create an Amazon Bedrock knowledge base to store the reference medical documents.
D. Configure an Amazon Bedrock knowledge base to store the reference medical document
E. Enable query decomposition in the knowledge bas
F. Configure an Amazon Bedrock flow that uses a foundation model and the knowledge base to support the application.
G. Use Amazon SageMaker AI to host custom ML models for both query decomposition and query expansio
H. Configure Amazon Bedrock knowledge bases to store the reference medical document
I. Encrypt the documents in the knowledge base.
J. Create an Amazon Bedrock agent to orchestrate multiple AWS Lambda functions to decompose querie
K. Create an Amazon Bedrock knowledge base to store the reference medical document
L. Use the agent??s built-in knowledge base capabilitie
M. Add deep research and reasoning capabilities to the agent to reduce ambiguity in the medical terminology.

Answer: B

NEW QUESTION 61
A medical device company wants to feed reports of medical procedures that used the company??s devices into an AI assistant. To protect patient privacy, the AI assistant must expose patient personally identifiable information (PII) only to surgeons. The AI assistant must redact PII for engineers. The AI assistant must reference only medical reports that are less than 3 years old.
The company stores reports in an Amazon S3 bucket as soon as each report is published. The company has already set up an Amazon Bedrock Knowledge Bases. The AI assistant uses Amazon Cognito to authenticate users.
Which solution will meet these requirements?

A. Enable Amazon Macie PII detection on the S3 bucke
B. Use an S3 trigger to invoke an AWS Lambda function that redacts PII from the report
C. Configure the Lambda function to delete outdated documents and invoke knowledge base syncing.
D. Invoke an AWS Lambda function to sync the S3 bucket and the knowledge base when a new report is uploade
E. Use a second Lambda function with Amazon Comprehend to redact PII for engineer
F. Use S3 Lifecycle rules to remove reports older than 3 years.
G. Set up an S3 Lifecycle configuration to remove reports that are older than 3 year
H. Schedule an AWS Lambda function to run daily syncs between the bucket and the knowledge bas
I. When users interact with the AI assistant, apply a guardrail configuration selected based on the user??s Cognito user group to redact PII from responses when required.
J. Create a second knowledge bas
K. Use Lambda and Amazon Comprehend to redact PII before syncing to the second knowledge bas
L. Route users to the appropriate knowledge base based on Cognito group membership.

Answer: C

NEW QUESTION 62
A financial services company is developing a real-time generative AI (GenAI) assistant to support human call center agents. The GenAI assistant must transcribe live customer speech, analyze context, and provide incremental suggestions to call center agents while a customer is still speaking. To preserve responsiveness, the GenAI assistant must maintain end-to-end latency under 1 second from speech to initial response display. The architecture must use only managed AWS services and must support bidirectional streaming to ensure that call center agents receive updates in real time.
Which solution will meet these requirements?

A. Use Amazon Transcribe streaming to transcribe call
B. Pass the text to Amazon Comprehend for sentiment analysi
C. Feed the results to Anthropic Claude on Amazon Bedrock by using the InvokeModel AP
D. Store results in Amazon DynamoD
E. Use a WebSocket API to display the results.
F. Use Amazon Transcribe streaming with partial results enabled to deliver fragments oftranscribed text before customers finish speakin
G. Forward text fragments to Amazon Bedrock by using the InvokeModelWithResponseStream AP
H. Stream responses to call center agents through an Amazon API Gateway WebSocket API.
I. Use Amazon Transcribe batch processing to convert calls to tex
J. Pass complete transcripts to Anthropic Claude on Amazon Bedrock by using the ConverseStream AP
K. Return responses through an Amazon Lex chatbot interface.
L. Use the Amazon Transcribe streaming API with an AWS Lambda function to transcribe each audio segmen
M. Call the Amazon Titan Embeddings model on Amazon Bedrock by using the InvokeModel AP
N. Publish results to Amazon SNS.

Answer: B

NEW QUESTION 63
A financial services company wants to develop an Amazon Bedrock application that gives analysts the ability to query quarterly earnings reports and financial statements. The financial documents are typically 5–100 pages long and contain both tabular data and text. The application must provide contextually accurate responses that preserve the relationship between financial metrics and their explanatory text. To support accurate and scalable retrieval, the application must incorporate document segmentation and context management strategies.
Which solution will meet these requirements?

A. Use a direct model invocation approach that uses Anthropic Claude to process each financial document as a single inpu
B. Use fine-tuned prompts that instruct the model to parse tables and text separately.
C. Use Amazon Bedrock Knowledge Bases to create a Retrieval Augmented Generation (RAG) application that retrieves relevant information from contextually chunked sections of financial document
D. Segment documents based on their structural layou
E. Include citations that reference the original source materials.
F. Deploy an Amazon Bedrock agent that has an action group that calls custom AWS Lambda functions to analyze financial document
G. Configure the Lambda functions to perform fixed-size chunking when a user submits a query about financial metrics.
H. Create one specialized Amazon Bedrock application that is optimized for structured dat
I. Create a second application that is optimized for unstructured dat
J. Configure each application to use a tailored chunking strategy that is suited to the application's content typ
K. Implement logic to link queries to the appropriate sources.

Answer: B

NEW QUESTION 64
A company is developing a generative AI (GenAI) application that analyzes customer service calls in real time and generates suggested responses for human customer service agents. The application must process 500,000 concurrent calls during peak hours with less than 200 ms end-to-end latency for each suggestion. The company uses existing architecture to transcribe customer call audio streams. The application must not exceed a predefined monthly compute budget and must maintain auto scaling capabilities.
Which solution will meet these requirements?

A. Deploy a large, complex reasoning model on Amazon Bedroc
B. Purchase provisioned throughput and optimize for batch processing.
C. Deploy a low-latency, real-time optimized model on Amazon Bedroc
D. Purchase provisioned throughput and set up automatic scaling policies.
E. Deploy a large language model (LLM) on an Amazon SageMaker real-time endpoint that uses dedicated GPU instances.
F. Deploy a mid-sized language model on an Amazon SageMaker serverless endpoint that is optimized for batch processing.

Answer: B

NEW QUESTION 67
A company uses an organization in AWS Organizations with all features enabled to manage multiple AWS accounts. Employees use Amazon Bedrock across multiple accounts. The company must prevent specific topics and proprietary information from being included in prompts to Amazon Bedrock models. The company must ensure that employees can use only approved Amazon Bedrock models. The company wants to manage these controls centrally.
Which combination of solutions will meet these requirements? (Select TWO.)

A. Create an IAM permissions boundary for each employee's IAM rol
B. Configure the permissions boundary to require an approved Amazon Bedrock guardrail identifier to invoke Amazon Bedrock model
C. Create an SCP that allows employees to use only approved models.
D. Create an SCP that allows employees to use only approved model
E. Configure the SCP to require employees to specify a guardrail identifier in calls to invoke an approved model.
F. Create an SCP that prevents an employee from invoking a model if a centrally deployed guardrail identifier is not specified in a call to the mode
G. Create a permissions boundary on each employee's IAM role that allows each employee to invoke only approved models.
H. Use AWS CloudFormation to create a custom Amazon Bedrock guardrail that has a block filtering polic
I. Use stack sets to deploy the guardrail to each account in the organization.
J. Use AWS CloudFormation to create a custom Amazon Bedrock guardrail that has a mask filtering polic
K. Use stack sets to deploy the guardrail to each account in the organization.

Answer: CD

NEW QUESTION 71
A company uses AWS Lambda functions to build an AI agent solution. A GenAI developer must set up a Model Context Protocol (MCP) server that accesses user information. The GenAI developer must also configure the AI agent to use the new MCP server. The GenAI developer must ensure that only authorized users can access the MCP server.
Which solution will meet these requirements?

A. Use a Lambda function to host the MCP serve
B. Grant the AI agent Lambda functions permission to invoke the Lambda function that hosts the MCP serve
C. Configure the AI agent??s MCP client to invoke the MCP server asynchronously.
D. Use a Lambda function to host the MCP serve
E. Grant the AI agent Lambda functions permission to invoke the Lambda function that hosts the MCP serve
F. Configure the AI agent to use the STDIO transport with the MCP server.
G. Use a Lambda function to host the MCP serve
H. Create an Amazon API Gateway HTTP API that proxies requests to the Lambda functio
I. Configure the AI agent solution to use the Streamable HTTP transport to make requests through the HTTP AP
J. Use Amazon Cognito to enforce OAuth 2.1.
K. Use a Lambda layer to host the MCP serve
L. Add the Lambda layer to the AI agent Lambda function
M. Configure the agentic AI solution to use the STDIO transport to send requests to the MCP serve

N. In the AI agent??s MCP configuration, specify the Lambda layer ARN as the comman
O. Specify the user credentials as environment variables.

Answer: C

NEW QUESTION 74
A hotel company wants to enhance a legacy Java-based property management system (PMS) by adding AI capabilities. The company wants to use Amazon Bedrock Knowledge Bases to provide staff with room availability information and hotel-specific details. The solution must maintain separate access controls for each hotel that the company manages. The solution must provide room availability information in near real time and must maintain consistent performance during peak usage periods.
Which solution will meet these requirements?

A. Deploy a single Amazon Bedrock knowledge base that contains combined data for all hotel
B. Configure AWS Lambda functions to synchronize data from each hotel??s PMS database through direct API connection
C. Implement AWS CloudTrail logging with hotel- specific filters to audit access logs for each hotel??s data.
D. Create an Amazon EventBridge rule for each hotel that is invoked by changes to the PMS databas
E. Configure the rule to send updates to a centralized Amazon Bedrock knowledge base in a management AWS accoun
F. Configure resource-based policies to enforce hotel-specific access controls.
G. Implement one Amazon Bedrock knowledge base for each hotel in a multi-account structur
H. Use direct data ingestion to provide near real-time room availability informatio
I. Schedule regular synchronization for less critical information.
J. Build a centralized Amazon Bedrock Agents solution that uses multiple knowledge base
K. Implement AWS IAM Identity Center with hotel-specific permission sets to control staff access.

Answer: C

NEW QUESTION 77
A company is implementing a serverless inference API by using AWS Lambda. The API will dynamically invoke multiple AI models hosted on Amazon Bedrock. The company needs to design a solution that can switch between model providers without modifying or redeploying Lambda code in real time. The design must include safe rollout of configuration changes and validation and rollback capabilities.
Which solution will meet these requirements?

A. Store the active model provider in AWS Systems Manager Parameter Stor
B. Configure a Lambda function to read the parameter at runtime to determine which model to invoke.
C. Store the active model provider in AWS AppConfi
D. Configure a Lambda function to read the configuration at runtime to determine which model to invoke.
E. Configure an Amazon API Gateway REST API to route requests to separate Lambda function
F. Hardcode each Lambda function to a specific model provide
G. Switch the integration target manually.
H. Store the active model provider in a JSON file hosted on Amazon S3. Use AWS AppConfig to reference the S3 file as a hosted configuration sourc
I. Configure a Lambda function to read the file through AppConfig at runtime to determine which model to invoke.

Answer: B

NEW QUESTION 82
A company is building a generative AI (GenAI) application that processes financial reports and provides summaries for analysts. The application must run two compute environments. In one environment, AWS Lambda functions must use the Python SDK to analyze reports on demand. In the second environment, Amazon EKS containers must use the JavaScript SDK to batch process multiple reports on a schedule. The application must maintain conversational context throughout multi-turn interactions, use the same foundation model (FM) across environments, and ensure consistent authentication.
Which solution will meet these requirements?

A. Use the Amazon Bedrock InvokeModel API with a separate authentication method for each environmen
B. Store conversation states in Amazon DynamoD
C. Use custom I/O formatting logic for each programming language.
D. Use the Amazon Bedrock Converse API directly in both environments with a common authentication mechanism that uses IAM role
E. Store conversation states in Amazon ElastiCach
F. Create programming language-specific wrappers for model parameters.
G. Create a centralized Amazon API Gateway REST API endpoint that handles all model interactions by using the InvokeModel AP
H. Store interaction history in application process memory in each Lambda function or EKS containe
I. Use environment variables to configure model parameters.
J. Use the Amazon Bedrock Converse API and IAM roles for authenticatio
K. Pass previous messages in the request messages array to maintain conversational contex
L. Use programming language-specific SDKs to establish consistent API interfaces.

Answer: D

NEW QUESTION 87
A company is creating a generative AI (GenAI) application that uses Amazon Bedrock foundation models (FMs). The application must use Microsoft Entra ID to authenticate. All FM API calls must stay on private network paths. Access to the application must be limited by department to specific model families. The company also needs a comprehensive audit trail of model interactions.
Which solution will meet these requirements?

A. Configure SAML federation between Microsoft Entra ID and AWS Identity and Access Managemen
B. Create department-specific IAM roles that allow only the required ModelId value
C. Create AWS PrivateLink interface VPC endpoints for Amazon Bedrock runtime service
D. Enable AWS CloudTrail to capture Amazon Bedrock API call
E. Configure Amazon Bedrock model invocation logging to record detailed model interactions.
F. Create an identity provider (IdP) connection in IAM to authenticate by using Microsoft Entra I
G. Assign department permission sets to control access to specific model familie

H. Deploy AWS Lambda functions in private subnets with a NAT gateway for egress to Amazon Bedrock public endpoint
I. Enable CloudWatch Logs to capture model interactions for auditing purposes.
J. Create a SAML identity provider (IdP) in IAM to authenticate by using Microsoft Entra I
K. Use IAM permissions boundaries to limit department roles' access to specific model familie
L. Configure public Amazon Bedrock API endpoints with VPC routing to maintain private network connectivit
M. Set up CloudTrail with Amazon S3 Lifecycle rules to manage audit logs of model interactions.
N. Configure OpenID Connect (OIDC) federation between Microsoft Entra ID and IA
O. Use attribute-based access control to map department attributes to specific model access permission
P. Apply SCP policies to restrict access to Amazon Bedrock FM families based on departmen
Q. Use Microsoft Entra ID's built-in logging capabilities to maintain an audit trail of model interactions.

Answer: A

NEW QUESTION 90
A financial services company is creating a Retrieval Augmented Generation (RAG) application that uses Amazon Bedrock to generate summaries of market activities. The application relies on a vector database that stores a small proprietary dataset with a low index count. The application must perform similarity searches. The Amazon Bedrock model??s responses must maximize accuracy and maintain high performance.
The company needs to configure the vector database and integrate it with the application. Which solution will meet these requirements?

A. Launch an Amazon MemoryDB cluster and configure the index by using the Flat algorith
B. Configure a horizontal scaling policy based on performance metrics.
C. Launch an Amazon MemoryDB cluster and configure the index by using the Hierarchical Navigable Small World (HNSW) algorith
D. Configure a vertical scaling policy based on performance metrics.
E. Launch an Amazon Aurora PostgreSQL cluster and configure the index by using the Inverted File with Flat Compression (IVFFlat) algorith
F. Configure the instance class to scale to a larger size when the load increases.
G. Launch an Amazon DocumentDB cluster that has an IVFFlat index and a high probe valu
H. Configure connections to the cluster as a replica se
I. Distribute reads to replica instances.

Answer: B

NEW QUESTION 92
An insurance company uses existing Amazon SageMaker AI infrastructure to support a web-based application that allows customers to predict what their insurance premiums will be. The company stores customer data that is used to train the SageMaker AI model in an Amazon S3 bucket. The dataset is growing rapidly. The company wants a solution to continuously re-train the model. The solution must automatically re-train and re-deploy the model to the application when an employee uploads a new customer data file to the S3 bucket.
Which solution will meet these requirements?

A. Use AWS Glue to run an ETL job on each uploaded fil
B. Configure the ETL job to use the AWS SDK to invoke the SageMaker AI model endpoin
C. Use real-time inference with the endpoint to re-deploy the model after it is re-trained on the updated customer dataset.
D. Create an AWS Lambda function and webhook handlers to generate an event when an employee uploads a new fil
E. Configure SageMaker Pipelines to re-deploy the model after it is re-trained on the updated customer datase
F. Use Amazon EventBridge to create an event bu
G. Set the Lambda function event as the source and SageMaker Pipelines as the target.
H. Create an AWS Step Functions Express workflow with AWS SDK integrations to retrieve the customer data from the S3 bucket when an employee uploads a new file to the S3 bucke
I. Use a SageMaker Data Wrangler flow to export the data from the S3 bucket to SageMaker Autopilo
J. Use the SageMaker Autopilot to re-deploy the model after it has been re-trained on the updated customer dataset.
K. Create an AWS Step Functions Standard workflo
L. Configure the first state to call an AWS Lambda function to respond when an employee uploads a new file to the S3 bucke
M. Use a pipeline in SageMaker Pipelines to re-deploy the model after it has been re-trained on the updated customer datase
N. Use the next state in the workflow to run the pipeline when the first state receives a response.

Answer: D

NEW QUESTION 93
A legal research company has a Retrieval Augmented Generation (RAG) application that uses Amazon Bedrock and Amazon OpenSearch Service. The application stores 768- dimensional vector embeddings for 15 million legal documents, including statutes, court rulings, and case summaries.
The company's current chunking strategy segments text into fixed-length blocks of 500 tokens. The current chunking strategy often splits contextually linked information such as legal arguments, court opinions, or statute references across separate chunks. Researchers report that generated outputs frequently omit key context or cite outdated legal information.
Recent application logs show a 40% increase in response times. The p95 latency metric exceeds 2 seconds. The company expects storage needs for the application to grow from 90 GB to 360 GB within a year.
The company needs a solution to improve retrieval relevance and system performance at scale.
Which solution will meet these requirements?

A. Increase the embedding vector dimensionality from 768 to 4,096 without changing the existing chunking or pre-processing strategy.
B. Replace dynamic retrieval with static, pre-written summaries that are stored in Amazon S3. Use Amazon CloudFront to serve the summaries to reduce compute demand and improve predictability.
C. Update the chunking strategy to use semantic boundaries such as complete legal arguments, clauses, or sections rather than fixed token limit
D. Regenerate vector embeddings to align with the new chunk structure.
E. Migrate from OpenSearch Service to Amazon DynamoD
F. Implement keyword-based indexes to enable faster lookups for legal concepts.

Answer: C

NEW QUESTION 97
A company is using Amazon Bedrock and Anthropic Claude 3 Haiku to develop an AI assistant. The AI assistant normally processes 10,000 requests each hour

but experiences surges of up to 30,000 requests each hour during peak usage periods. The AI assistant must respond within 2 seconds while operating across multiple AWS Regions.
The company observes that during peak usage periods, the AI assistant experiences throughput bottlenecks that cause increased latency and occasional request timeouts. The company must resolve the performance issues.
Which solution will meet this requirement?

A. Purchase provisioned throughput and sufficient model units (MUs) in a single Regio
B. Configure the application to retry failed requests with exponential backoff.
C. Implement token batching to reduce API overhea
D. Use cross-Region inference profiles to automatically distribute traffic across available Regions.
E. Set up auto scaling AWS Lambda functions in each Regio
F. Implement client-side round-robin request distributio
G. Purchase one model unit (MU) of provisioned throughput as a backup.
H. Implement batch inference for all requests by using Amazon S3 buckets across multiple Region
I. Use Amazon SQS to set up an asynchronous retrieval process.

Answer: B

NEW QUESTION 102
A company provides a service that helps users from around the world discover new restaurants. The service has 50 million monthly active users. The company wants to implement a semantic search solution across a database that contains 20 million restaurants and 200 million reviews. The company currently stores the data in a PostgreSQL database.
The solution must support complex natural language queries and return results for at least 95% of queries within 500 ms. The solution must maintain data freshness for restaurant details that update hourly. The solution must also scale cost-effectively during peak usage periods.
Which solution will meet these requirements with the LEAST development effort?

A. Migrate the restaurant data to Amazon OpenSearch Servic
B. Implement keyword-based search rules that use custom analyzers and relevance tuning to find restaurants based on attributes such as cuisine type, feature, and locatio
C. Create Amazon API Gateway HTTP API endpoints to transform user queries into structured search parameters.
D. Migrate the restaurant data to Amazon OpenSearch Servic
E. Use a foundation model (FM) in Amazon Bedrock to generate vector embeddings from restaurant descriptions, reviews, and menu item
F. When users submit natural language queries, convert the queries to embeddings by using the same F
G. Perform k-nearest neighbors (k-NN) searches to find semantically similar results.
H. Keep the restaurant data in PostgreSQL and implement a pgvector extensio
I. Use a foundation model (FM) in Amazon Bedrock to generate vector embeddings from restaurant dat
J. Store the vector embeddings directly in PostgreSQ
K. Create an AWS Lambda function to convert natural language queries to vector representations by using the same F
L. Configure the Lambda function to perform similarity searches within the database.
M. Migrate the restaurant data to an Amazon Bedrock knowledge base by using a custom ingestion pipelin
N. Configure the knowledge base to automatically generate embeddings from restaurant informatio
O. Use the Amazon Bedrock Retrieve API with built-in vector search capabilities to query the knowledge base directly by using natural language input.

Answer: D

NEW QUESTION 104
A company is developing a customer support application that uses Amazon Bedrock foundation models (FMs) to provide real-time AI assistance to the company??s employees. The application must display AI-generated responses character by character as the responses are generated. The application needs to support thousands of concurrent users with minimal latency. The responses typically take 15 to 45 seconds to finish.
Which solution will meet these requirements?

A. Configure an Amazon API Gateway WebSocket API with an AWS Lambda integratio
B. Configure the WebSocket API to invoke the Amazon Bedrock InvokeModelWithResponseStream API and stream partial responses through WebSocket connections.
C. Configure an Amazon API Gateway REST API with an AWS Lambda integratio
D. Configure the REST API to invoke the Amazon Bedrock standard InvokeModel API and implement frontend client-side polling every 100 ms for complete response chunks.
E. Implement direct frontend client connections to Amazon Bedrock by using IAM user credentials and the InvokeModelWithResponseStream API without any intermediate gateway or proxy layer.
F. Configure an Amazon API Gateway HTTP API with an AWS Lambda integratio
G. Configure the HTTP API to cache complete responses in an Amazon DynamoDB table and serve the responses through multiple paginated GET requests to frontend clients.

Answer: A

NEW QUESTION 107
A media company must use Amazon Bedrock to implement a robust governance process for AI-generated content. The company needs to manage hundreds of prompt templates. Multiple teams use the templates across multiple AWS Regions to generate content. The solution must provide version control with approval workflows that include notifications for pending reviews. The solution must also provide detailed audit trails that document prompt activities and consistent prompt parameterization to enforce quality standards.
Which solution will meet these requirements?

A. Configure Amazon Bedrock Studio prompt template
B. Use Amazon CloudWatch dashboards to display prompt usage metric
C. Store approval status in Amazon DynamoD
D. Use AWS Lambda functions to enforce approvals.
E. Use Amazon Bedrock Prompt Management to implement version contro
F. Configure AWS CloudTrail for audit loggin
G. Use AWS Identity and Access Management policies to control approval permission
H. Create parameterized prompt templates by specifying variables.

I. Use AWS Step Functions to create an approval workflo
J. Store prompts in Amazon S3. Use tags to implement version contro
K. Use Amazon EventBridge to send notifications.
L. Deploy Amazon SageMaker Canvas with prompt templates stored in Amazon S3. Use AWS CloudFormation for version contro
M. Use AWS Config to enforce approval policies.

Answer: B

NEW QUESTION 110
A financial services company is developing a customer service AI assistant application that uses a foundation model (FM) in Amazon Bedrock. The application must provide transparent responses by documenting reasoning and by citing sources that are used for Retrieval Augmented Generation (RAG). The application must capture comprehensive audit trails for all responses to users. The application must be able to serve up to 10,000 concurrent users and must respond to each customer inquiry within 2 seconds.
Which solution will meet these requirements with the LEAST operational overhead?

A. Enable tracing for Amazon Bedrock Agent
B. Configure structured prompts that direct the FM to provide evidence presentation
C. Integrate Amazon Bedrock Knowledge Bases with data sources to enable RA
D. Configure the application to reference and cite authoritative conten
E. Deploy the application in a Multi-AZ architectur
F. Use Amazon API Gateway and AWS Lambda functions to scale the applicatio
G. Use Amazon CloudFront to provide low- latency delivery.
H. Enable tracing for Amazon Bedrock agent
I. Integrate a custom RAG pipeline with Amazon OpenSearch Service to retrieve and cite source
J. Configure structured prompts to present retrieved evidenc
K. Deploy the application behind an Amazon API Gateway REST AP
L. Use AWS Lambda functions and Amazon CloudFront to scale the application and to provide low latenc
M. Store logs in Amazon S3 and use AWS CloudTrail to capture audit trails.
N. Use Amazon CloudWatch to monitor latency and error rate
O. Embed model prompts directly in the application backend to cite source
P. Store application interactions with users in Amazon RDS for audits.
Q. Store generated responses and supporting evidence in an Amazon S3 bucke
R. Enable versioning on the bucket for audit
S. Use AWS Glue to catalog retrieved document
T. Process the retrieved documents in Amazon Athena to generate periodic compliance reports.

Answer: A

NEW QUESTION 114
A GenAI developer is evaluating Amazon Bedrock foundation models (FMs) to enhance a Europe-based company's internal business application. The company has a multi-account landing zone in AWS Control Tower. The company uses Service Control Policies (SCPs) to allow its accounts to use only the eu-north-1 and eu-west-1 Regions. All customer data must remain in private networks within the approved AWS Regions.
The GenAI developer selects an FM based on analysis and testing and hosts the model in the eu-central-1 Region and the eu-west-3 Region. The GenAI developer must enable access to the FM for the company's employees. The GenAI developer must ensure that requests to the FM are private and remain within the same Regions as the FM.
Which solution will meet these requirements?

A. Deploy an AWS Lambda function that is exposed by a private Amazon API Gateway REST API to a VPC in eu-north-1. Create a VPC endpoint for the selected FM in eu- central-1 and eu-west-3. Extend existing SCPs to allow employees to use the F
B. Integrate the REST API with the business application.
C. Deploy the FM on Amazon EC2 instances in eu-north-1. Deploy a private Amazon API Gateway REST API in front of the EC2 instance
D. Configure an Amazon Bedrock VPC endpoin
E. Integrate the REST API with the business application.
F. Configure the FM to use cross-Region inference through a Europe-scoped endpoin
G. Configure an Amazon Bedrock VPC endpoin
H. Extend existing SCPs to allow employees to use the FM through inference profiles in Europe-based Regions where the FM is availabl
I. Use an inference profile to integrate Amazon Bedrock with the business application.
J. Deploy the FM in Amazon SageMaker in eu-north-1. Configure a SageMaker VPC endpoin
K. Extend existing SCPs to allow employees to use the SageMaker endpoin
L. Integrate the FM in SageMaker with the business application.

Answer: C

NEW QUESTION 118
A company is developing a generative AI (GenAI) application by using Amazon Bedrock. The application will analyze patterns and relationships in the company??s data. The application will process millions of new data points daily across AWS Regions in Europe, North America, and Asia before storing the data in Amazon S3. The application must comply with local data protection and storage regulations. Data residency and processing must occur within the same continent. The application must also maintain audit trails of the application??s decision-making processes and provide data classification capabilities.
Which solution will meet these requirements?

A. Deploy the application in each Region with local IAM policie
B. Use Amazon Bedrock cross-Region inference to distribute the workloa
C. Use Amazon CloudWatch to log AI decision-making processe
D. Manually track compliance certifications across Regions.
E. Use SCPs with AWS Organizations to manage location-specific permission
F. Use AWS CloudTrail immutable logs to audit decision-making processe
G. Import a custom model into Amazon Bedrock and deploy the model to each Region.
H. Use Amazon S3 Object Lock with Region-specific S3 bucket policie
I. Pre-process the data points within the Region based on geographic origin before sending the data points to Amazon Bedroc
J. Use Amazon Macie to classify the dat

K. Use AWS CloudTrail immutable logs to audit the decision-making processes.
L. Create separate AWS accounts for each Region with individual compliance framework
M. Use Amazon SageMaker AI with custom monitorin
N. Create manual compliance reports for each regulatory jurisdiction.

Answer: C

NEW QUESTION 123
A company is developing a customer communication platform that uses an AI assistant powered by an Amazon Bedrock foundation model (FM). The AI assistant summarizes customer messages and generates initial response drafts.
The company wants to use Amazon Comprehend to implement layered content filtering. The layered content filtering must prevent sharing of offensive content, protect customer privacy, and detect potential inappropriate advice solicitation. Inappropriate advice solicitation includes requests for unethical practices, harmful activities, or manipulative behaviors.
The solution must maintain acceptable overall response times, so all pre-processing filters must finish before the content reaches the FM.
Which solution will meet these requirements?

A. Use parallel processing with asynchronous API call
B. Use toxicity detection for offensive conten
C. Use prompt safety classification for inappropriate advice solicitatio
D. Use personally identifiable information (PII) detection without redaction.
E. Use custom classification to build an FM that detects offensive content and inappropriate advice solicitatio
F. Apply personally identifiable information (PII) detection as a secondary filter only when messages pass the custom classifier.
G. Deploy a multi-stage proces
H. Configure the process to use prompt safety classification first, then toxicity detection on safe prompts only, and finally personally identifiable information (PII) detection in streaming mod
I. Route flagged messages through Amazon EventBridge for human review.
J. Use toxicity detection with thresholds configured to 0.5 for all categorie
K. Use parallel processing for both prompt safety classification and personally identifiable information (PII) detection with entity redactio
L. Apply Amazon CloudWatch alarms to filter metrics.

Answer: D

NEW QUESTION 124
A company is designing a solution that uses foundation models (FMs) to support multiple AI workloads. Some FMs must be invoked on demand and in real time. Other FMs require consistent high-throughput access for batch processing.
The solution must support hybrid deployment patterns and run workloads across cloud infrastructure and on-premises infrastructure to comply with data residency and compliance requirements.
Which combination of steps will meet these requirements? (Select TWO.)

A. Use AWS Lambda to orchestrate low-latency FM inference by invoking FMs hosted on Amazon SageMaker AI asynchronous endpoints.
B. Configure provisioned throughput in Amazon Bedrock to ensure consistent performance for high-volume workloads.
C. Deploy FMs to Amazon SageMaker AI endpoints with support for edge deployment by using Amazon SageMaker Ne
D. Orchestrate the FMs by using AWS Lambda to support hybrid deployment.
E. Use Amazon Bedrock with auto-scaling to handle unpredictable traffic surges.
F. Use Amazon SageMaker JumpStart to host and invoke the FMs.

Answer: BC

NEW QUESTION 127
.....

Thank You for Trying Our Product

* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

* One year free update

You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

We currently serve more than 30,000,000 customers.

* Shop Securely

All transactions are protected by VeriSign!

100% Pass Your AIP-C01 Exam with Our Prep Materials Via below:

https://www.certleader.com/AIP-C01-dumps.html