SAS A00-402

SAS MACHINE LEARNING CERTIFICATION QUESTIONS & ANSWERS

Exam Summary – Syllabus – Questions

A00-402

SAS Certified Specialist - Machine Learning Using SAS Viya 3.5
50-55 Questions Exam – 65% Cut Score – Duration of 100 minutes

Table of Contents

Know Your A00-402 Certification Well:

The A00-402 is best suitable for candidates who want to gain knowledge in the SAS Advanced Analytics. Before you start your A00-402 preparation you may struggle to get all the crucial SAS Machine Learning materials like A00-402 syllabus, sample questions, study guide.

But don't worry the A00-402 PDF is here to help you prepare in a stress free manner.

The PDF is a combination of all your queries like-

- What is in the A00-402 syllabus?
- How many questions are there in the A00-402 exam?
- Which Practice test would help me to pass the A00-402 exam at the first attempt?

Passing the A00-402 exam makes you SAS Certified Specialist - Machine Learning Using SAS Viya 3.5. Having the SAS Machine Learning certification opens multiple opportunities for you. You can grab a new job, get a higher salary or simply get recognition within your current organization.

A00-402 SAS Machine Learning Certification Details:

Exam Name	SAS Certified Specialist - Machine Learning Using SAS Viya 3.5
Exam Code	A00-402
Exam Duration	100 minutes
Exam Questions	50-55
Passing Score	65%
Exam Price	$180 (USD)
Training	**Machine Learning Using SAS Viya**
Books	**Machine Learning with SAS® Viya**
Exam Registration	**Pearson VUE**
Sample Questions	**SAS Machine Learning Certification Sample Question**
Practice Exam	**SAS Machine Learning Certification Practice Exam**

A00-402 Syllabus:

Objective	Details
Data Sources (30%)	
Create a project in Model Studio	- Bring data into Model Studio for analysis • Import data from a local source (Import tab) • Add data from a stored data source (Data Sources tab) • Use an in-memory data source (Available tab) - Create Model Studio Pipelines with the New Pipeline window • Automatically generate pipelines • Pipeline templates - Advanced Advisor options • Maximum class level • Maximum % missing • Interval cut-off - Partition data into training, validation, and test • Explain why partitioning is important • Explain the different methods to partition data (stratified vs simple random) - Use Event Based Sampling to oversample for rare events.
Explore the data	- Use the DATA EXPLORATION node - Profile data during data definition - Preliminary data exploration using the data tab - Save data with the SAVE DATA node
Modify data	- Modify metadata with the MANAGE VARIABLES node - Use the REPLACEMENT node to update variable values - Use the TRANSFORMATION node to correct problems with input data sources, such as variables distribution or outliers - Use the IMPUTE node to impute missing values and create missing value indicators - Modify data within the DATA tab
Reduce the dimensionality of the data	- Use the FEATURE EXTRACTION node - Prepare text data for modeling with the TEXT MINING node
Use the VARIABLE SELECTION node to identify important variables to be	- Unsupervised Selection - Fast Supervised Selection - Linear Regression Selection - Decision Tree Selection - Forest Selection

Objective	Details
included in a predictive model	- Gradient Boosting Selection - Create Validation from Training - Use multiple methods within the same VARIABLE SELECTION node.
Building Models (50%)	
Describe key supervised machine learning terms and concepts	- Data partitioning: training, validation, test data sets - Observations (cases), independent (input) variables/features, dependent (target) variables - Measurement scales: Interval, ordinal, nominal (categorical), binary variables - Prediction types: decisions, rankings, estimates - Dimensionality, redundancy, irrelevancy - Decision trees, neural networks, regression models - Model optimization, overfitting, underfitting, model selection - Describe ensemble models
Build models with decision trees and ensemble of trees	- Explain how decision trees identify split points • Split search algorithm • Recursive partitioning • Decision tree algorithms • Multiway vs. binary splits • Impurity reduction • Gini, entropy, Bonferroni, IRG, FTEST, variance • Compare methods to grow decision trees for categorical vs continuous response variables - Explain the effect of missing values on decision trees - Explain surrogate rules - Explain the purpose of pruning decision trees - Explain bagging vs. boosting methods - Build models with the DECISION TREE node • Adjust splitting options • Adjust pruning options - Build models with the GRADIENT BOOSTING node • Adjust general options: number of trees, learning rate, L1/L2 regularization rate • Adjust Tree Splitting options • Adjust early stopping • Adjust autotuning - Build models with the FOREST node • Adjust number of trees • Adjust tree splitting options

Objective	Details
	• Adjust autotuning - Interpret decision tree, gradient boosting, and forest results (fit statistics, output, tree diagrams, tree maps, variable importance, error plots, autotuned results)
Build models with neural networks	- Describe the characteristics of neural network models • Adaptive learning • Universal approximation • Neurons, hidden layers, perceptrons, multilayer perceptrons • Weights and bias • Activation functions • Optimization Methods (LBFGS and Stochastic Gradient Descent) • Variable standardization - Build models with the NEURAL NETWORK node • Adjust number of layers and neurons • Adjust optimization options and early stopping criterion - Interpret NEURAL NETWORK node results (network diagram, iteration plots, and output)
Build models with support vector machines	- Describe the characteristics of support vector machines. - Build model with the SVM node • Adjust general properties (Kernel, Penalty, Tolerance) • Perform Autotuning - Interpret SVM node results (Output)
Use Model Interpretability tools to explain black box models	- Partial Dependence plots - Individual Conditional Expectation plots - Local Interpretable Model-Agnostic Explanations plots - Kernel-SHAP plots
Incorporate externally written code	- Open Source Code node - SAS Code node - Score Code Import node
Model Assessment and Deployment (20%)	
Explain the principles of Model Assessment	- Explain different dimensions for model comparison • Training speed • Model application speed • Tolerance

Objective	Details
	• Model clarity - Explain honest assessment • Evaluate a model with a holdout data set - Use the appropriate fit statistic for different prediction types • Average error for estimates • Misclassification for decisions
Assess and compare models in Model Studio	- Compare models with the MODEL COMPARISON node - Compare models with the PIPELINE COMPARISON tab - Interpret Fit Statistics, Lift Reports, ROC reports.
Deploy a model	- Exporting score code - Registering a model - Publish a model

SAS A00-402 Sample Questions:

Question: 1

Which feature extraction method can take both interval variables and class variables as inputs?

a) Principal component analysis
b) Autoencoder
c) Singular value decomposition
d) Robust PCA

Answer: b

Question: 2

As the number of input variables in a problem increases, there is an exponential increase in the number of observations needed to densely populate the feature space. This is referred to as:

a) Problem of rare events
b) Multicollinearity
c) Curse of Dimensionality
d) Underfitting

Answer: c

Question: 3

What is another term for a feature in predictive modeling?

a) Instance
b) Input
c) Target
d) Outcome

Answer: b

Question: 4

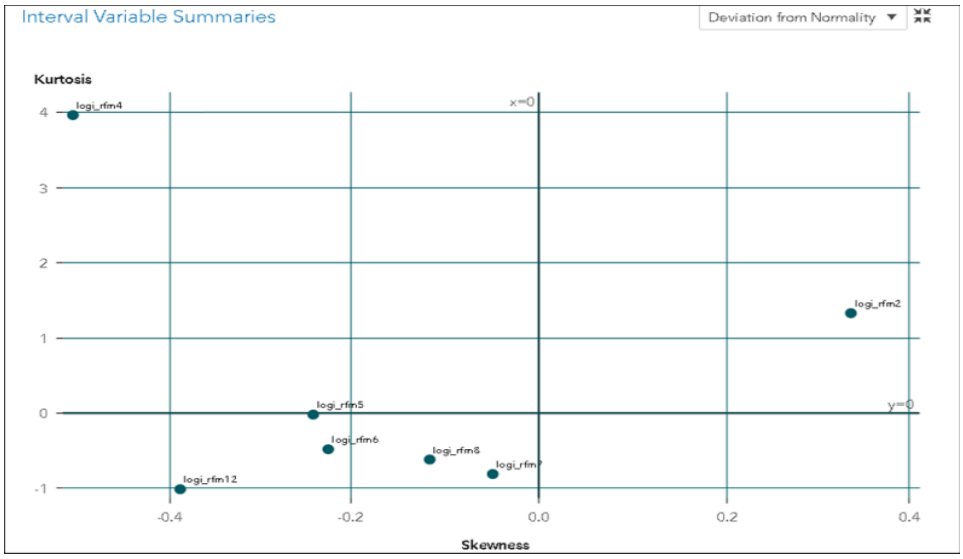
A project has been created and a pipeline has been run in Model Studio. Which project setting can you edit?

a) Advisor Options for missing values
b) Rules for model comparison statistic
c) Partition Data percentages
d) Event-based Sampling proportions

Answer: b

Question: 5

Refer to the exhibit below



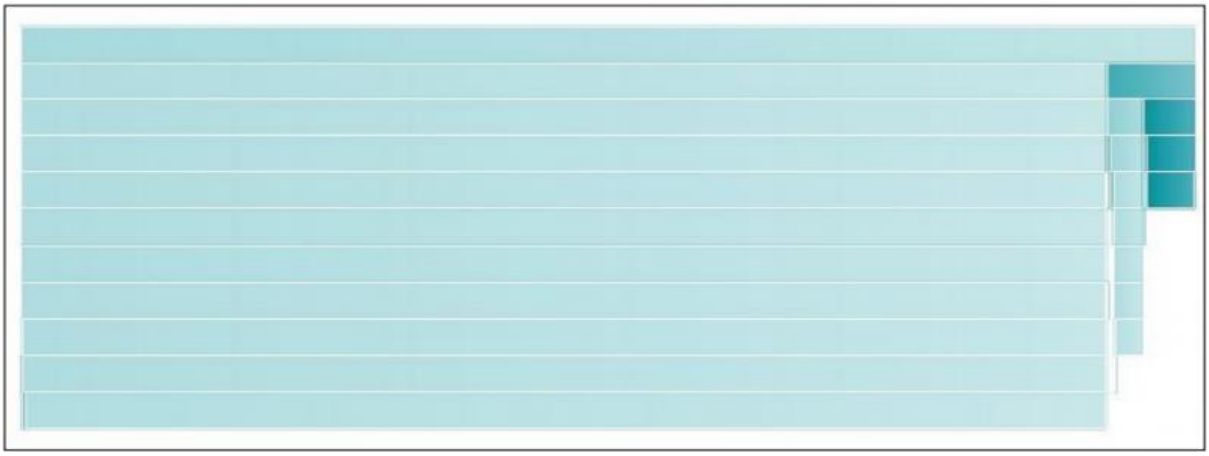
Based on the output from the Data Exploration node shown in the exhibit, which variable has the most thin tails (most platykurtic distribution)?

a) Logi_rfm4
b) Logi_rfm6
c) Logi_rfm8
d) Logi_rfm12

Answer: d

Question: 6

Refer to the treemap shown in the exhibit below



Which statement is true about the tree map for a decision tree with a binary target?

a) The top bar represents the node with the highest probability of event
b) The darker bars represent nodes with a lower probability of event.
c) The top bar represents the node with the highest count
d) The wider bars represent nodes with a higher probability of event.

Answer: c

Question: 7

Which statement is true regarding decision trees and models based on ensembles of trees?

a) In the gradient boosting algorithm, for all but the first iteration, the target is the residual from the previous decision tree model.
b) For a Forest model, the out-of-bag sample is simply the original validation data set from when the raw data partitioning took place.
c) In the Forest algorithm, each individual tree is pruned based on using minimum Average Squared Error.
d) A single decision tree will always be outperformed by a model based on an ensemble of trees.

Answer: a

Question: 8

In Model Studio, you have multiple pipelines in a project. Which statement is true?

a) The Model Comparison node compares only the champion models for each project.
b) The Pipeline Comparison tab compares all of the models from each pipeline.
c) You can override the champion in a Model Comparison node.
d) You can override the champion in a Pipeline Comparison tab.

Answer: d

Question: 9

Given the following properties for a neural network model, which statement is true regrading hidden units in the model? The following SAS program is submitted:


Properties

Property name	Property value
missAsLevl	false
inputStd	STD
nHidden	1
hiddenAll	false
hiddenAllNum	50
actFuncAll	TANH
hidden1	26
actFunc1	TANH

a) There are no hidden units in the model
b) The number of hidden units is 1.
c) The number of hidden units is 50.
d) The number of hidden units is 26.

Answer: d

Question: 10

Which statements are true for the F1 score?

(Choose 2.)

a) F1 score is calculated based on a depth value
b) F1 score is calculated based on a cut off value
c) F1 score is applicable to a model with a binary target.
d) F1 score is applicable to a model with an interval target.

Answer: b, c

Study Guide to Crack SAS Machine Learning A00-402 Exam:

- Getting details of the A00-402 syllabus, is the first step of a study plan. This pdf is going to be of ultimate help. Completion of the syllabus is must to pass the A00-402 exam.
- Making a schedule is vital. A structured method of preparation leads to success. A candidate must plan his schedule and follow it rigorously to attain success.
- Joining the SAS provided training for A00-402 exam could be of much help. If there is specific training for the exam, you can discover it from the link above.
- Read from the A00-402 sample questions to gain your idea about the actual exam questions. In this PDF useful sample questions are provided to make your exam preparation easy.
- Practicing on A00-402 practice tests is must. Continuous practice will make you an expert in all syllabus areas.