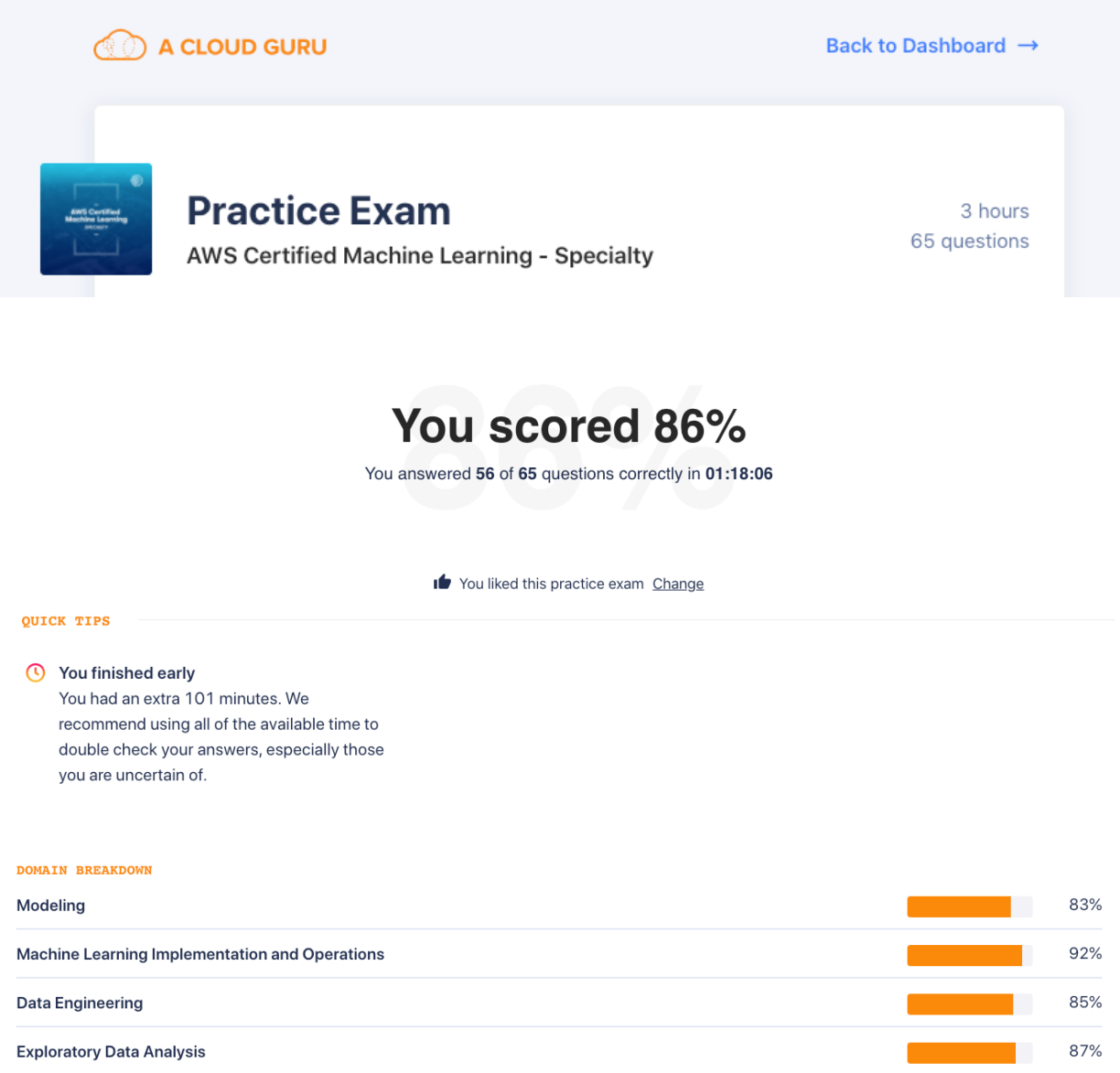
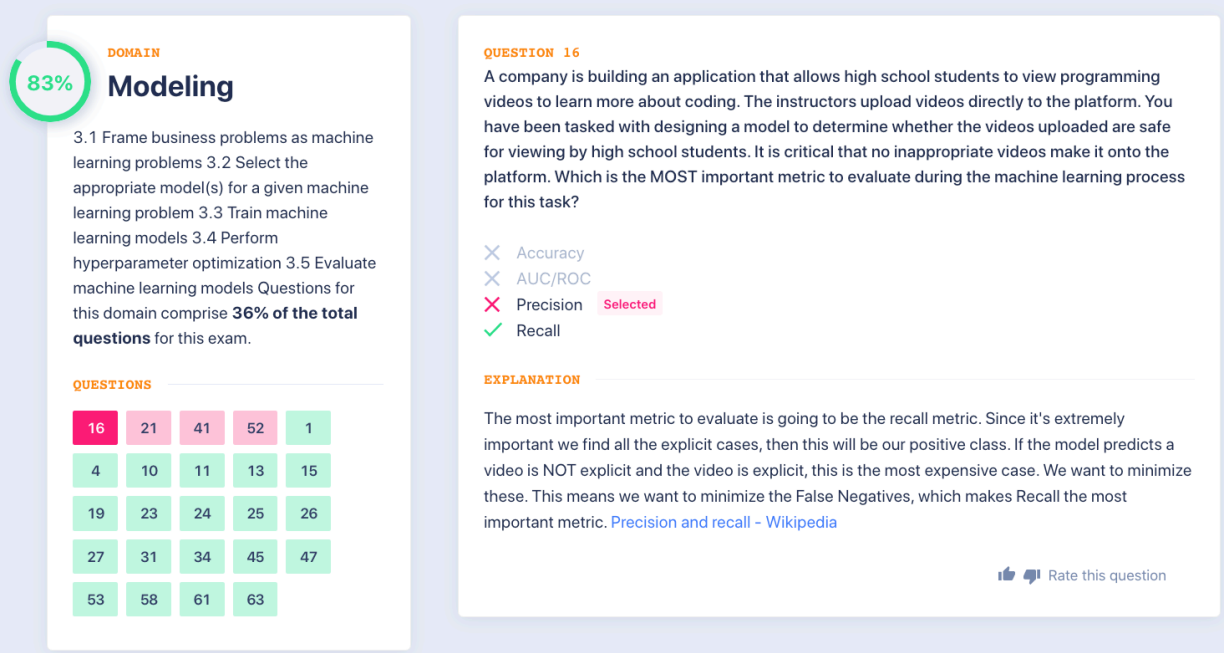
CloudGuru - Practice Exam - 3rd attempt

https://practice-exam.acloud.guru/aws-certified-machine-learning-specialty?courseId=aws-certified-machine-learning-specialty&_ga=2.151769904.849115380.1605014235-906779002.1602959981&_gac=1.251718395.1602959981.CjwKCAjwrKr8BRB_EiwA7eFaphu95ArJm-GaAsgiegqwqcDAUIPEYAwg5tU6E1-93RcNJU2Y23-VXRoCzacQAvD_BwE



Questions I did not get right:



I put the question is: Is video safe? Y/N as the question refers to “safe”.
We want to limit false positives, where we say a video is safe but it is not => Precision

But the other way to ask the question I: IS the video explicit? Y/N
In that case, we want to limit false negatives, where we way a video is not explicit but it is => Recall

So it all comes down the the formulation of the binary classifier problem

QUESTION 21

You are consulting with a large educational organization on a ML model using the built-in BlazingText SageMaker algorithm. They have asked for help deciding which metric to use in an automatic model tuning job. What can you recommend to help them get started?

- ✓ The metrics to use for optimization can vary depending on how you are using the algorithm.
- ✗ Metrics with a prefix of validation: are the ones to always use for optimization jobs. Selected
- ✗ The proper metric for BlazingText optimization is validation:accuracy.
- ✗ Metrics with a prefix of train: are the ones to always use for optimization jobs.
- ✗ BlazingText is not supported with hyperparameter tuning.

EXPLANATION

BlazingText optimization is an example of where we would use different metrics depending on how we are using the algorithm. You should consult the documentation for the algorithm you are using to determine the proper metric. Tune a BlazingText Model - Amazon SageMaker

I agree with that as I knew we could use BlazingText in Word2Vec mode or Text Classification.
But what I selected is correct too.

QUESTION 41

You are trying to follow an old family recipe for making sourdough bread. The recipe states that you should add enough water to the flour such that it is slightly sticky. Which if the following best characterizes this instruction?

- ✗ Algorithm Selected
- ✗ Supervised Learning
- ✓ Heuristic
- ✗ Unsupervised Learning
- ✗ Reinforcement Learning
- ✗ Metric

EXPLANATION

A heuristic technique, often called simply a heuristic, is any approach to problem solving or self-discovery that employs a practical method, not guaranteed to be optimal, perfect, or rational, but instead sufficient for reaching an immediate goal. Where finding an optimal solution is impossible or impractical, heuristic methods can be used to speed up the process of finding a satisfactory solution. A heuristic is an educated guess or intuition that usually does not ensure an optimal outcome. By asking the chef to evaluate the quantity of water needed based on individual perception best embodies a heuristic. ALGORITHMS AND HEURISTICS

QUESTION 52

A binary classification model has been created to sort parts on an assembly line into acceptable or unacceptable, based on a complex array of readings. The model incorrectly decides that some flawed parts are acceptable when they should have been marked as unacceptable. Which of the following correctly defines this type of result?

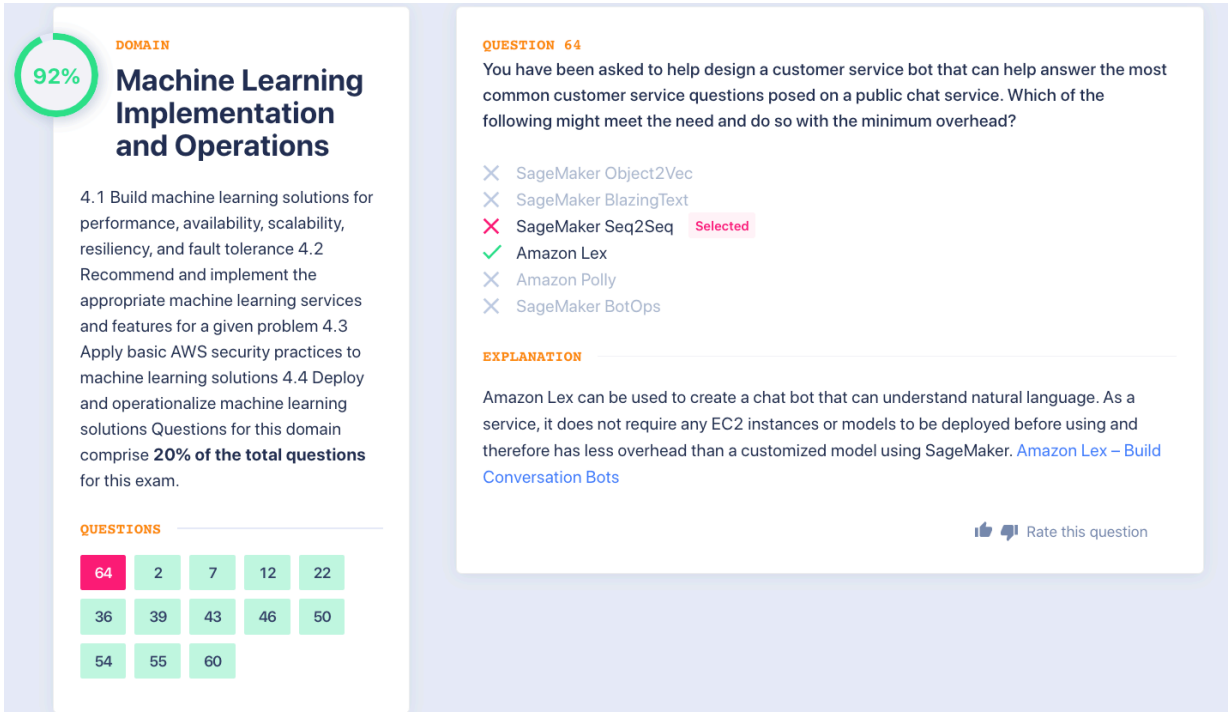
- ✓ False Negative
- ✗ False Positive Selected
- ✗ Type I Error Selected
- ✗ True Negative
- ✓ Type II Error
- ✗ True Positive

EXPLANATION

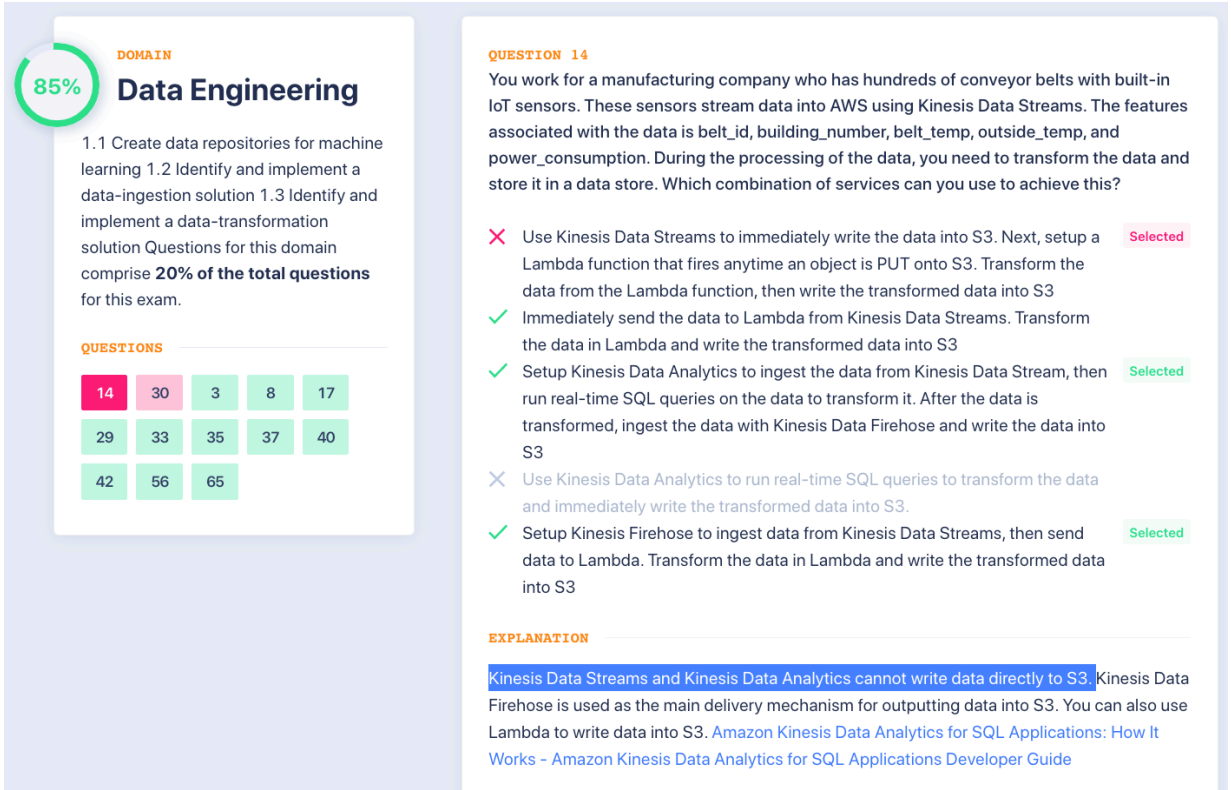
A false negative is when a model does not properly flag something that is should have. In this case, the flawed parts that got through as good parts were incorrectly identified as ok when they were bad. A false negative is also known as a Type II error. Type I and type II errors - Wikipedia

Same here, formatting the question is key.
It should be: is a part “unacceptable”?
In that case Type II makes sense with False negatives (not this part is acceptable when it is not)

I had formulated as: is a part “acceptable?



That one was obvious - I misread it and should have focused on “bot”



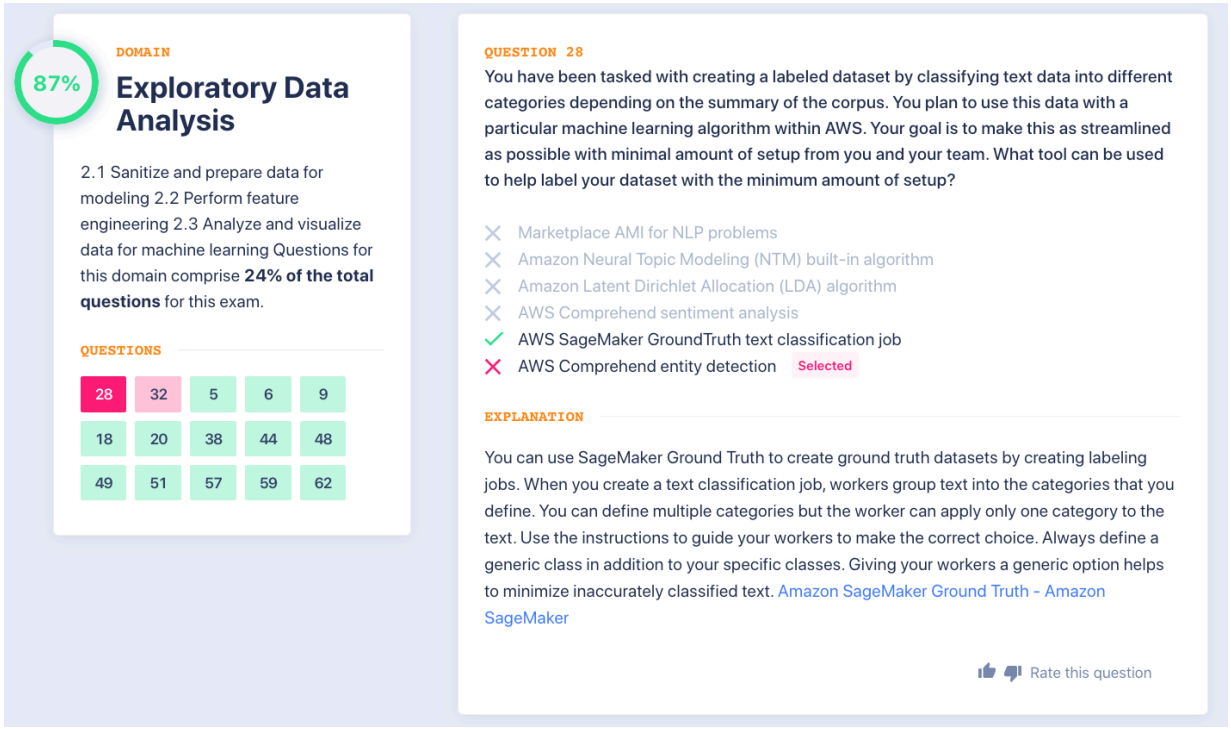
QUESTION 30

A machine learning model is being created using Amazon's Factorization Machines algorithm to help make click predictions and item recommendations for new customers. Which of the following would be candidates during the training process?

- ✓ Creating a binary classification model where the testing dataset is scored using Binary Cross Entropy (Log Loss), Accuracy, and F1 Score.
- ✗ Using sparse data in CSV format as training data.
- ✗ Making inferences to the model in application/csv format.
- ✓ Using sparse data in recordIO-protobuf format with Float32 tensors as training data. Selected
- ✓ Creating a regression model where the testing dataset is scored using Root Mean Square Error (RMSE). Selected
- ✗ Creating a multi-classification model where the testing dataset is scored using Area Under The Curve (AUC). Selected

EXPLANATION

The factorization machine algorithm can be run in either in binary classification mode or regression mode. In regression mode, the testing dataset is scored using Root Mean Square Error (RMSE). In binary classification mode, the test dataset is scored using Binary Cross Entropy (Log Loss), Accuracy (at threshold=0.5) and F1 Score (at threshold =0.5). For training, the factorization machines algorithm currently supports only the recordIO-protobuf format with Float32 tensors. CSV format is not a good candidate. For inference, factorization machines support the application/json and x-recordio-protobuf formats. Factorization Machines Algorithm - Amazon SageMaker



I hesitated on that one between the 2