Module 3 Quiz

1. Suppose we are building a recommendation system for an online retailer of home improvement products. Which of the below options would be the best outcome metric to use to evaluate the success of our project?
    1. Mean squared error
    2. Dollars saved on inventory holding costs
    3. Recall
    4. Average additional sales dollars earned per customer visit to the site

2. Outcome metrics are commonly stated in terms of (select all that apply):
    - Dollars earned or saved
    - Error rate
    - Area under curve
    - Time saved

3. When do we typically use model output metrics (select all that apply)?
    - To compare models and perform model selection
    - To evaluate the final model prior to deployment
    - To state the success criteria for solving the user problem we are trying to address
    - To monitor the ongoing performance of deployed models

4. What is one key difference between Mean Squared Error (MSE) and Mean Absolute Error (MAE)?
    1. MSE is a regression metric and MAE is a classification metric
    2. MAE is much more commonly used than MSE
    3. The magnitude of MSE is influenced by the scale of your data, while MAE is always the same magnitude regarless of the magnitude of values of your input features
    4. MSE is more heavily influenced by outliers than MAE because it penalizes large errors heavily due to the squared term in the calculation

5. Why is Mean Absolute Percentage Error (MAPE) popular among non-technical audiences relative to MSE and MAE for quanitfying the error of regression models?
    1. It can be more intuitively understood than MSE and MAE in terms of relative error, since it normalizes the scale of the error to the scale of the values to be predicted (e.g. MAPE = 6%)
    2. It provide a more precise quantification of error
    3. It penalizes large errors more than MAE and MSE
    4. It communicates error as an absolute number rather than a relative number

6. Suppose I am building a classification model for a university to try to identify students likely to have Covid-19 (positives) vs. no Covid-19 (negatives) based on daily questionaires about symptoms, and my main goal is to make sure I identify everyone with Covid-19 correctly in order to keep them from entering campus. I calculate the recall (or true postive rate) of my model as 92%. What does this number mean?

1. Out of all students my model predicted as being likely to have Covid-19, 92% of them actually had Covid-19
2. Out of all the predictions my model made (positive or negative), 92% of them were correct
3. Out of all the students who tested negative for Covid-19, my model incorrectly classified 92% of them as likely positive
4. Out of all the students who were in fact postive for Covid-19, my model correctly identified 92% of them as being likely positive

7. Which of the following is an incorrect statement regarding Receiver Operating Characteristic (ROC) curves?

1. ROC curves show the tradeoff between false positive rate and true positive rate for a model
2. The AUROC (area under ROC curve) for a model ranges from -1 to +1
3. A perfect model would have an AUROC (area under ROC curve) of 1
4. The points along the ROC curve represent the combination of true positive rate and false positive rate for different threshold values of the model

8. I am building a model for an insurance company to predict which insured drivers will likely have a car accident within the next year. I calculate the accuracy of my model as 97.8%, and proclaim success. Why might my declaration of success be misleading? Keep in mind that only a small fraction of insured drivers have a car accident within a given year.

1. I should have used a regression metric in this situation rather than accuracy, a classification metric
2. An accuracy of 97.8% is poor for this problem
3. My model might be predicting "no accident" for every insured driver, and 97.8% of the time it would be right, resulting in a high accuracy score even though the model itself was worthless. I would be better of using a different metric instead of accuracy to evaluate my model
4. My model is exceptional and I am correct to declare victory

9. I am building a binary classification (positive/negative) model and use the typical threshold value of 0.5, meaning that if the output probability is above 0.5 the model predicts a positive, and if it below 0.5 the model predicts a negative. What would happen to the recall of my model if I reduced my threshold to 0.3?

1. Recall would go down because I would classify less observations as positive and therefore would correctly identify less of the positive cases
2. Recall would go up because I would classify more observations as positives and therefore would correctly identify more of the positive cases
3. Recall would not change because it does not depend on the threshold value
4. Recall would go up because more of the predicted positives were actually positives

10. What is the value of displaying the confusion matrix for a classification model?
    1. It shows us which classes the model is correctly predicting and which classes it is struggling to correctly predict
    2. It shows us the distribution of actual target values among the classes we are trying to predict, but does not give us any information about the distribution of predictions from our model
    3. It allows us to easily calculate the Area under ROC curve (AUROC)
    4. It helps us visualize the outputs of different models on the same chart so that we can compare performance between models