Logistic Regression, Regularization, & ROC Curves - Interview Question Solutions

1. What is the difference between a parameter and a hyperparameter?

- A model parameter describes the final model itself, e.g. slope in a linear model.

- A learning hyperparameter describes the way in which a model parameter is learned,

e.g. learning rate, penalty terms, number of features to include in a weak predictor.

2. Name and briefly explain several evaluation metrics that are used for classification.

- Accuracy - measures the percentage of the time you correctly classified samples:

(true positive + true negative) / all samples

- Precision - measures the percentage of the predicted positives that were

correctly classified: true positives / (true positives + false positives)

- Sensitivity or Recall - measures the percentage of true positives that were

correctly classified: true positives / (true positives + false negative)

- Specificity - measures the percentage of true negatives that were

correctly classified: true negatives / (true negatives + false positives)

- F1 - measurement that balances accuracy and precision

(or you can think of it as balancing Type I and Type II error)

- AUC

3. What are two common ways to automate hyperparameter tuning?

- Grid search - test every possible combination of pre-defined hyperparameter

values and select the best one

- Randomized search - randomly test possible combinations of pre-defined

hyperparameter values and select the best tested one

4. When should you use no regularization vs ridge vs lasso vs elastic net?

- Regularized models are useful to prevent overfitting or to reduce variance.

- Lasso can be effective when you want to use to automatically do

feature selection in order to create a simpler model but can be dangerous

since it may be erratic and remove features that contain useful signal.

- Elastic net is a balance of ridge and lasso, and it can be used to the

same effect as lasso with less erratic behavior.

5. Is logistic regression a regressor or a classifier?

- Logistic regression is usually used as a classifier because it predicts

discrete classes.

- Having said that, it technically outputs a continuous value associated

with each prediction. So we see that is is actually a regression algorithm,

hence the name, that can solve classification problems. It is fair to say that

it is a classifier because it is used for classification,

although it is also technically also a regressor.

6. What parameters be tuned in logistic regression models? Explain how they affect model learning.

- Logistic regression models can be tuned using regularizations techniques

(commonly L2 norm, but other norms may be used as well)

7. What is regularization? Why would you use it?

- Regularization adds a penalty to the loss function which prevents the model

from fitting too closely to the training data.

- This is used when the model is overfit/has high variance to reduce the variance.

8. How does logistic regression work?

- Logistic regression creates a linear model with the log odds of the target.

- It predicts a value between 0 and 1, which is the predicted probability of

being in the positive class.

9. How is the ROC curve created?

- The ROC curve plots the true positive rate vs. the false positive rate

for every classification threshold.

10. How would you interpret a coefficient from a logistic regression model?

- “For every one-unit increase in [X feature], the log odds that the

observation is in [positive y class] are [coefficient] times as

large as the odds that the observation is not in [positive y class],

holding all else constant.”

- Note: by exponentiating the coefficient, we can interpret this in terms

of odds rather than log odds.