Python Lab Update (10 points)

If you did not use a gradient boosted decision tree or random forest for the Python Lab to predict length of stay, you will have to create a GBDT or random forest for this assignment. Turn in your code in either .pdf or .html format for this assignment. Answers to the questions can be included in the code.

1. Take your best model and re-tune it using hyperopt or hyperband. Retrain your model with the best parameters. After retraining, answer the following questions:
    - What is your AUC on training and testing? (2 points)
    - Is your model overfit, underfit, or fit well? (2 points)
    - If your model is overfit or underfit, retune your model to improve its performance (manually or with gridsearch). Describe the approach you took to improve your model (1 point).
    - If you already used hyperband or hyperopt, please discuss your tuning with these methods and resubmit the same code. You can additionally try a different hyperparameter tuning method if you'd like.
2. Take your best model and find your variable importance using LIME or SHAP. Plot your variable importance summaries for LIME or SHAP on both your training and testing data (2 points). Then answer the following questions:
    - What are your most important variables? (1 point)
    - How do they affect the model as a whole? Or How do they affect a certain observation? (1 point)