Name: _____

In-class final, EC524/424

150 points possible

Instructions **Briefly** answer the following questions/prompts. Typically, 1–3 short (and complete) sentences will suffice. We will deduct points for excessively long answers.

1. (7.5 points) Briefly explain how *prediction* fundamentally differs from *causal inference*.

2. (7.5 points) What is the difference between supervised and unsupervised machine learning?

3. (7.5 points) How does logistic regression differ from OLS regression?

4. (7.5 points) What is an ensemble?

5. (7.5 points) Why do we normalize predictors in penalized models?

6. Let's talk about the bias-variance tradeoff.

   (a) (7.5 points) What do we mean by *bias* (in the bias-variance tradeoff)?

   (b) (7.5 points) What do we mean by *variance* (in the bias-variance tradeoff)?

   (c) (7.5 points) Explain the bias-variance tradeoff.

7. (7.5 points) How do random forests *decorrelate* their trees?

8. (7.5 points) Why do we typically prefer k-fold cross-validation to the validation-set approach?

9. (7.5 points) Why might we care more about *precision* than *accuracy*?

10. (7.5 points) Define the null classifier. Then explain the null classifier helps us understand a classifier's performance.

11. (7.5 points) Explain how the L1 and L2 norms differ and why this difference is meaningful.

12. (7.5 points) Why do classification trees often split using entropy or GINI rather than accuracy?

13. (7.5 points) Imagine you have three machine-learning models: elastic net, ridge regression, and the lasso. Which of the models implicitly performs variable selection?

14. (7.5 points) Define "imputation" and explain why it is important to prediction problems.

15. (7.5 points) Describe *how and why* support vector machines expand the predictor space?

16. (7.5 points) Which methods help avoid overfitting with machine-learning models?

17. (7.5 points) What is bootstrapping, and how is it useful?

18. (7.5 points) Sketch out a confusion matrix (it doesn't need numbers) and explain how it helps understand a model's predictions.