

ADS 503 - Applied Predictive Modeling (M5)

Summer 2024 - Week 5

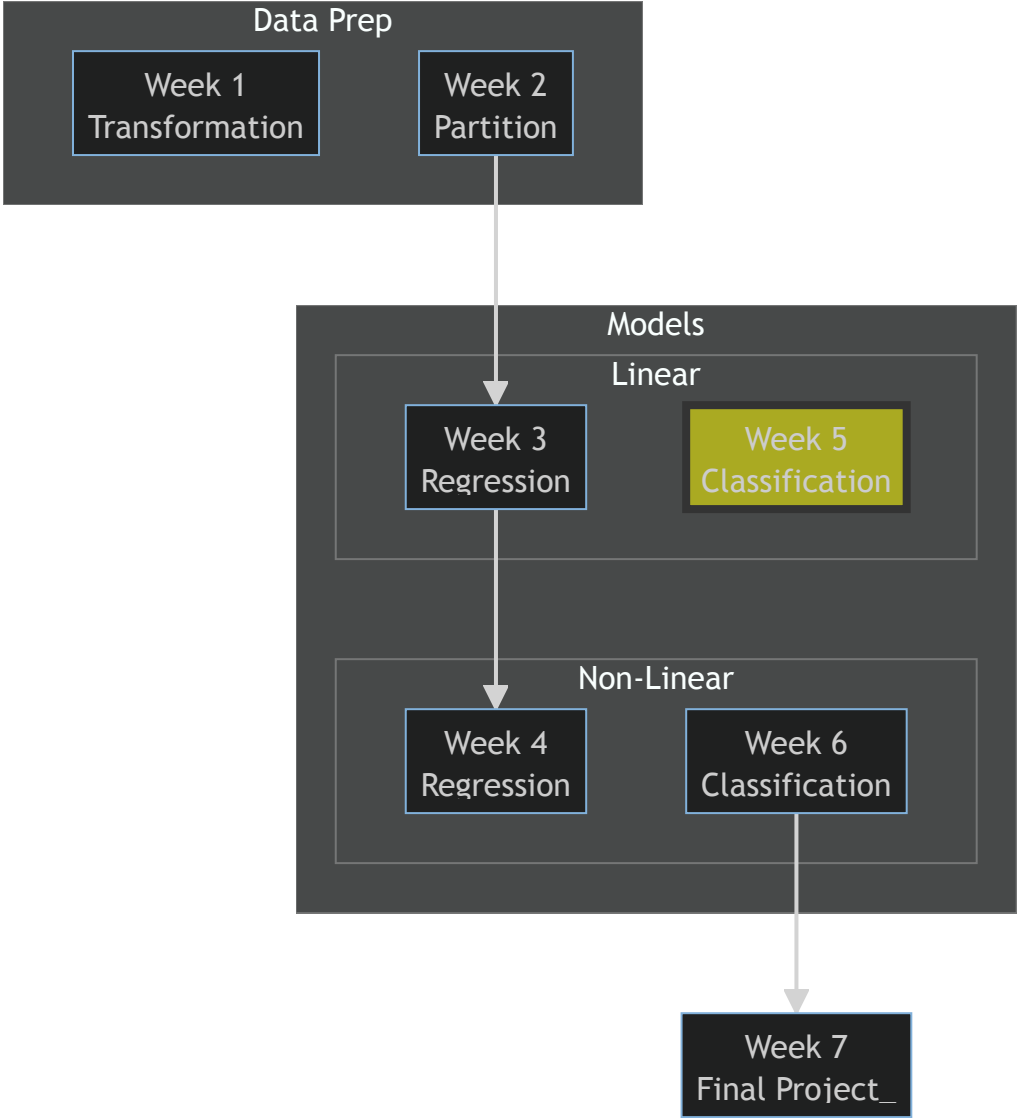
Dave Hurst

Start Recording!

Agenda

- Posit.Cloud - note: monthly billing cycle rollover
- Course Map
- Assignment 4 Review
- Assignment 5 Tips
- QA

Course Map



Assignment 4 Review

We are baking here!



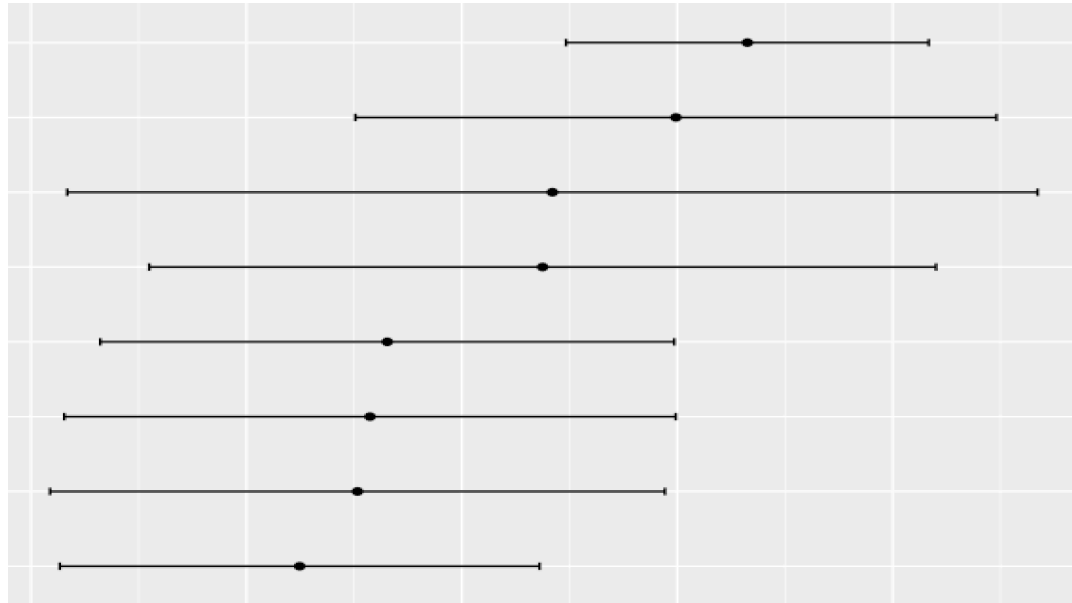
Ref: tidymodels.org

Assignment 5 Tips

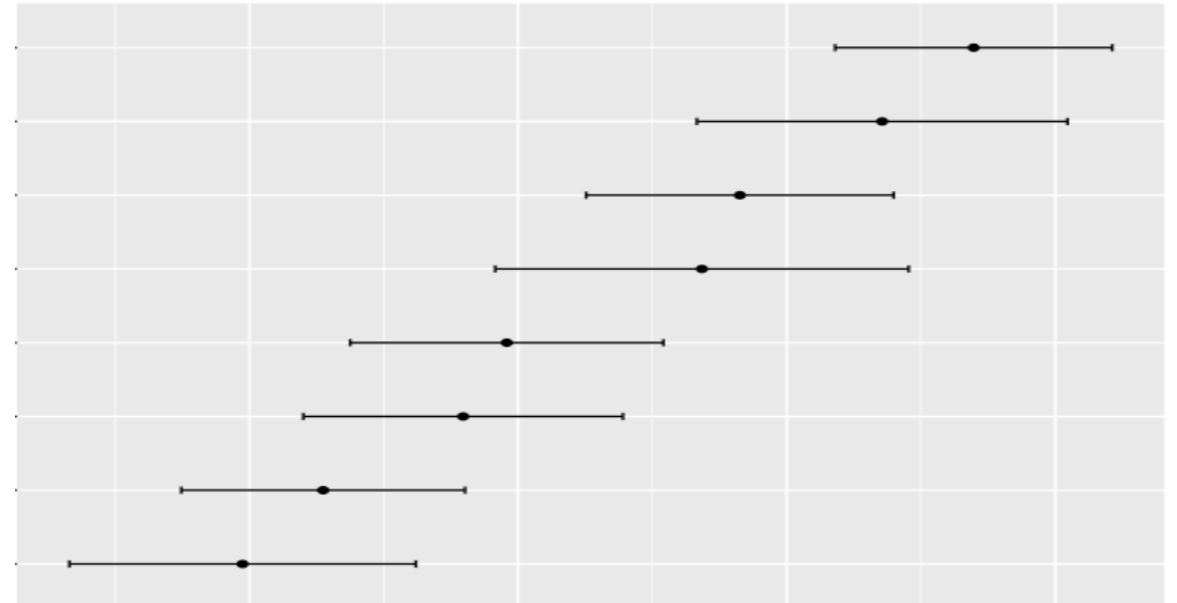
- 5.1 ROC scores are low
- 5.2 Accuracy scores are high
- remember the imbalanced class exercise from Assignment 2
- plot tuning (where appropriate)
- consider confidence intervals when evaluating models
- try to shrink CI (how?)

A5 - Shrinking CI

```
1 ctrl <- trainControl(method = "repeatedcv", repeats =  
2                       summaryFunction = twoClassSum  
3                       classProbs = TRUE,  
4                       savePredictions = TRUE)
```



```
1 ctrl <- trainControl(method = "repeatedcv", repeats =  
2                       summaryFunction = twoClassSum  
3                       classProbs = TRUE,  
4                       savePredictions = TRUE)
```



- Total Notebook runtime ~ 3 mins on posit.cloud
- equivalent to `method = "cv"`

- (model names and time intentionally omitted)

Note: `summaryFunction = [twoClassSummary | multiClassSummary]`

A5 - Useful Code Patterns

Importance Ranking

```
1 model_imp <- varImp(best_full_model)
2 model_top5 <- model_imp$importance |>
3   rownames_to_column(var = 'Predictor') |>
4   arrange(-Overall) |>
5   head(5) |>
6   select(ends_with('Predictor'), Importance = Overall)
```

Accuracy Summary Tables

```
1 extract_accuracy <- function(model = "model", cm) {
2   #cm from confusionMatrix(pred, reference = truth)
3   tibble(
4     model = model,
5     Accuracy = cm$overall['Accuracy'],
6     `CI Lower` = cm$overall['AccuracyLower'],
7     `CI Upper` = cm$overall['AccuracyUpper']
8   )
9 }
10
11 accuracies <- bind_rows(
12   extract_accuracy('LDA', oil_lda_cm),
13   extract_accuracy('PLR', oil_plr_cm),
14   extract_accuracy('NSC', oil_nsc_cm)
15 )
16
17 accuracies |>
18   arrange(-Accuracy) |>
19   gt() |>
20   fmt_number(decimals = 3)
```


Q&A