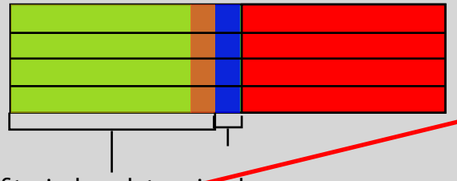
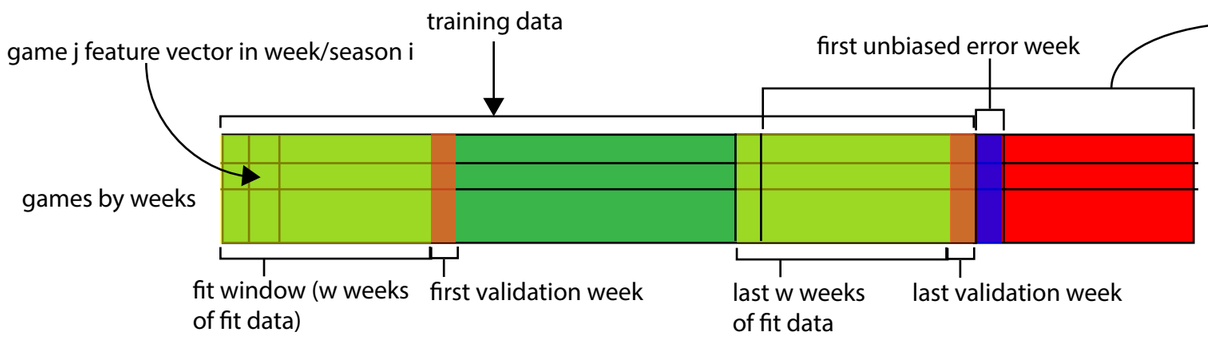


fit window determined
in stage 1 validation

1. For each week in test data, follow best dynamic system to generate a prediction rule.
2. Evaluate mean bet prediction accuracy
3. Average over each week for final prediction accuracy

?????

1. Choose best row (dynamic system for generating a model for a week of games) from level 2 merged validation results
2. Use the system chosen to evaluate approximate unbiased test error using the held out test data



model blending; repeat model selection using level 2 validation
? should the choice of u be made first for each base model?

For each base model (i.e., boosted trees, k-NN), perform 1 and 2

1 For w in fit window sizes:
- a. For m in 1:number of model configurations:
  - i. For v in 1:(n-w):
    1. Fit model with data in columns v:(v+w-1)
    2. Evaluate model summary statistic/s by making predictions with data in column v+w
    3. Store summary statistics in validation results table for w in cell (m,v)

2 For u in validation window sizes:
- a. For x in 1:(n-w-u):
  - i. Choose best row, model configuration, given validation window summ stats summary statistics (orange)
  - ii. Use best model (model corresponding to the row with the bes summary statistics) to generate predictions for final prediction on week x
  - ii. Store predictions or summary statistics in level 2 validadtion results table for u in element x