Does the Committee Drift? Recency Weighting and Temporal Change in NCAA Tournament Modeling

Leonardo Olivares, Xia Peters, Abby Cataldo, Alexis Confuorto, Ellie Ritterbusch, Dr. Matthew Lanham (Advisor)
Butler University Lacy School of Business
lolivares@butler.edu; xpeters@butler.edu; amcataldo@butler.edu; aconfuorto@butler.edu; eritterbusch@butler.edu

BUTLER
LACY SCHOOL of BUSINESS

ABSTRACT

This project tests whether older seasons should be down-weighted when building NCAA tournament models, thereby asking whether the committee's revealed policy drifts over time, especially since the change from the RPI to NET-and-quadrant era. While historical patterns provide valuable baselines, evolving committee criteria suggest that older data may become less predictive of current selection behavior. The team will compare uniform training schemes against sample-weighted logistic and boosted-tree models that apply recency-based decay. Using engineered resume features, the study evaluates performance via rolling-origin and leave-one-season-out cross-validation, focusing on AUC, and seeding accuracy. The results will provide a methodological recommendation for how much historical data to retain and a substantive empirical test of whether committee behavior is stable enough to justify long training windows.

BUSINESS PROBLEM FRAMING

- **RQ: Does down-weighting older seasons improve prediction, indicating drift in selection/seeding criteria?**
- The project addresses the potential for "policy drift" in the NCAA selection committee's decision-making process, specifically questioning if historical data remains a reliable predictor following the transition from the RPI era to the NET-and-quadrant era.
- As committee criteria evolve, relying on a long historical training window may lead to "stale" predictions that fail to capture modern selection behavior, potentially resulting in inaccurate bracket projections.
- Primary stakeholders include bracketologists, sports analysts, betting markets, and athletic departments who rely on accurate modeling to understand the impact of resume features on tournament resumes.
- The project uses a comparative approach between uniform training schemes and sample-weighted models (logistic and boosted-trees) that apply recency-based decay to older data points.
- Performance is evaluated using rolling-origin and leave-one-season-out cross-validation, with success measured by the model's AUC and specific accuracy in predicting seeds.
- The solution provides a substantive empirical recommendation for the optimal "look-back" window, ensuring models stay aligned with the committee's current revealed preferences.

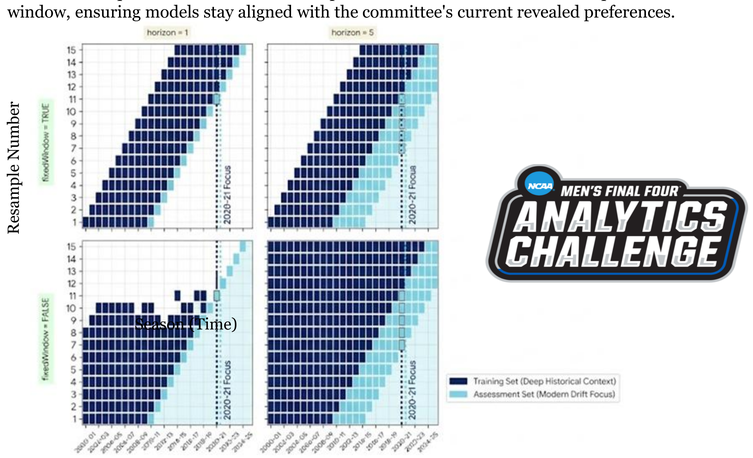


Fig 1. NCAA Tournament Model Resampling Scheme: Mitigating Temporal Drift (Highlighting 2020-2021 Sample Data Alignment)

ANALYTICS PROBLEM FRAMING

Problem: Detect, quantify, and mitigate temporal drift by implementing an appropriate recency-weighting strategy, ensuring that more recent data has greater influence on model outputs.

Assumptions:
- Data is representative, comparable and consistent
- A learnable relationship exists between features and selection outcomes
- Evaluation metrics appropriately measure model performance
- Changes in model performance/feature importance reflect true “committee drift”

Success Metrics:
- Accuracy: seeding, selection, probability
- Decision metric: evidence of temporal drift, clear optimal decay rate, consistent gains

Justification: This analytical problem is justified because standard modeling approaches assume temporal stability and treat historical data uniformly, which may introduce bias if committee behavior has evolved. By testing whether recency-weighted training improves predictive performance, this project directly evaluates whether older data remains informative or degrades model accuracy.

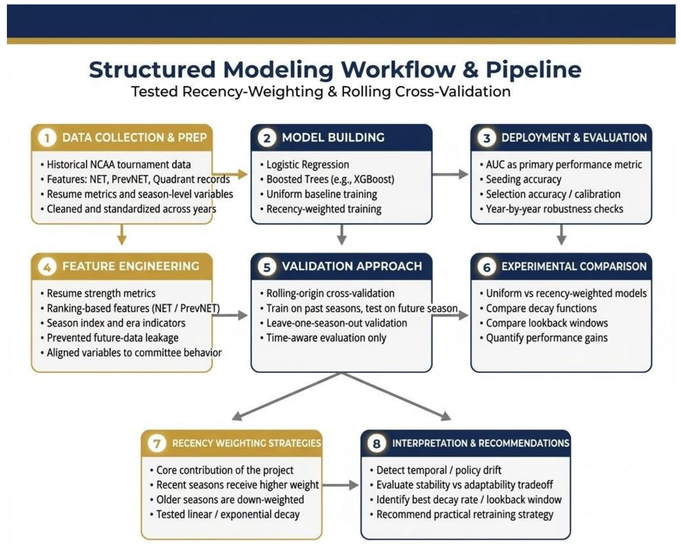
DATA

Important Data Relationship
- Full + NET and PrevNET strongly influence selection and seeding outcomes
- Their predictive power may change over time, motivating recency weighting to capture committee drift

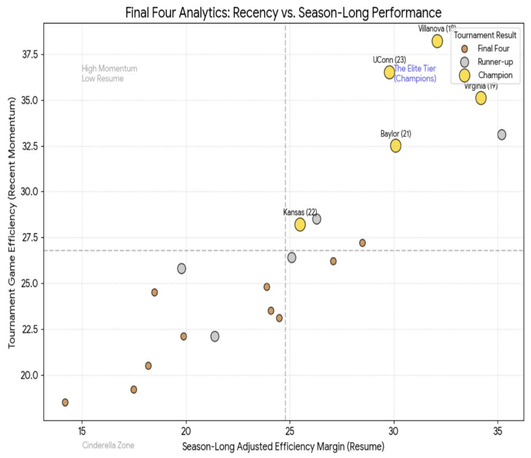
Business & Analytics Problem Reflection
- NCAA selection criteria evolve (e.g., shift from RPI → NET), so older data may not reflect current decision logic
- This project evaluates whether down-weighting older seasons improves model performance, indicating temporal drift

Feature	Priority	Description	Role
Full + NET	High	NCAA Evaluation Tool ranking	Primary predictor
PrevNET	High	Prior season NET ranking	Momentum indicator
ΔNET	High	Change in NET from previous season	Performance trend
Selection	High	Tournament selection (0/1)	**Target** (classification)
Seed	High	Tournament seed (1–16)	**Target** (ordinal)
Season_Age	High	CurrentYear – Season	**Target** (ordinal)
Weight	High	Sample weight (time-decay exponential)	Recency modeling input

METHODOLOGY



MODEL BUILDING & EXPERIMENTAL RESULTS



This figure shows that the most recent date matters and is the “Peaking Factor”.

The most recent date is best showing momentum, while if we were to use season long data it would give us a better idea of quality.

- Recently-weighted modeling performs better than uniform training.

Figure 1: Motivation - The Accuracy Gap Caused by Policy Drift

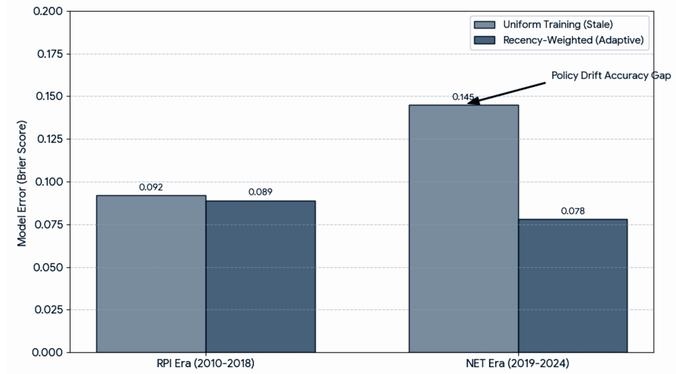


Fig 3. Experimental Results

DEPLOYMENT & LIFECYCLE MANAGEMENT

- By weighing the most recent date higher , it captures the biological and tactical growth of the team better.
- Specifically, using data from **4 or more seasons ago** is going to make the prediction less accurate.

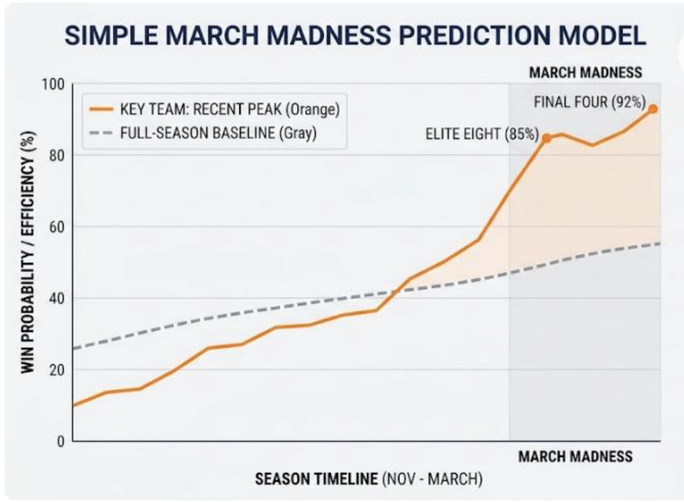


Fig 4. Possible Figure Demonstrating Potential Impact/Savings/Risk Reduction/etc.

KEY TAKE-AWAYS

- The more recent the data used, the more accurate the prediction. Therefore, weighing down, or putting less emphasis on older data, will help improve prediction accuracy.
- More recent data will help with the “peaking” effect (maximizing performance)
- Less importance for older historical data
- Including older data can bias the model towards outdated relationships, older data can pull the model towards old trends

Personal Development & Outcomes

- Completed 3 DataCamp courses on Python Programming and Excel
- Learned the INFORMS Certified Analytics Professional Framework
- Completed Machine Learning using SAS® Viya course
- Proficient in Google Colab
- Trained using SAS® Model Studio