Learning-to-Rank Models for Tournament Seeding Prediction Using Pairwise Ranking and Season-Based Validation

Gabrielle Joslin, Alex Kanipe, Madelline De La Torre-Castillo, Caroline Bosh, Maximous Popp, Dr. Matthew Lanham (Advisor)
Butler University Lacy School of Business
gjoslin@butler.edu; akanipe@butler.edu; mdelatorre@butler.edu; cbosh@butler.edu, mpopp@butler.edu

BUTLER
LACY SCHOOL *of* BUSINESS

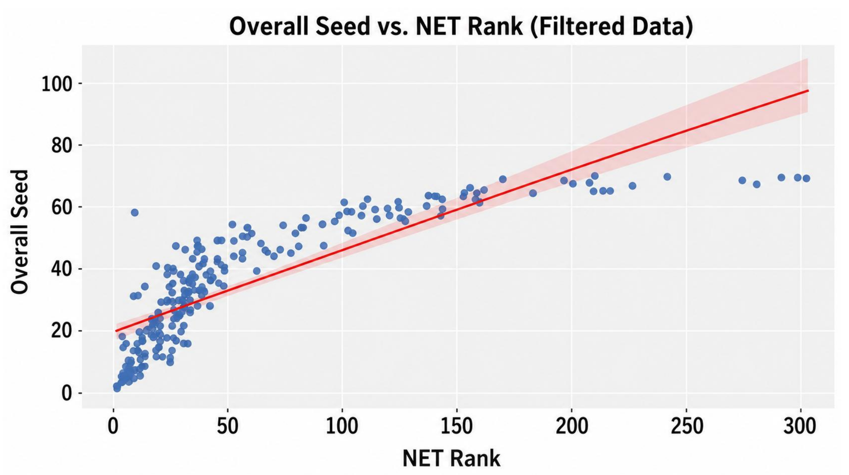
ABSTRACT

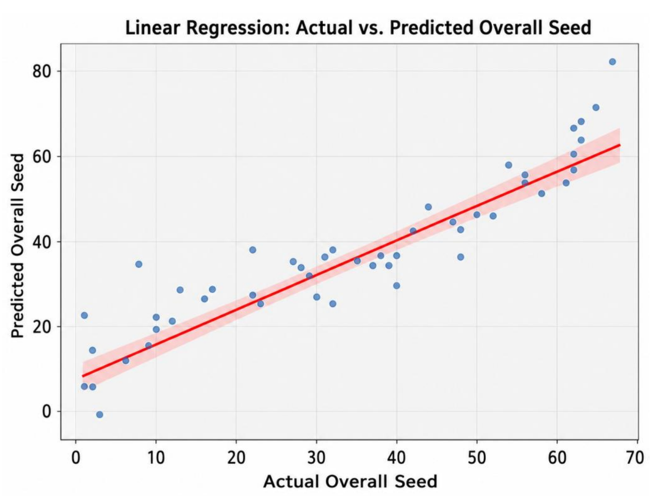
This project proposes a ranking-first view of NCAA tournament forecasting by asking whether within-season learning-to-rank methods can improve both field identification and seed assignment. The idea is closely aligned with recent selection-method research. Feldman's Powerwise framework emphasizes transparent pairwise comparisons built from head-to-head results, common opponents, and rating information, while policy-capturing work by Reinig and Horowitz shows how committee rankings can be represented as an underlying preference structure even when exact weights are difficult to recover. Inspired by that literature, we construct pairwise team comparisons within season and estimate ranking models that produce an ordered list rather than isolated in-or-out predictions. The resulting ranking will be mapped into the tournament field through a top-N rule and into seed lines through rank-to-quartet conversion. Evaluation will include rank agreement, missed bids, and seed-line accuracy under leave-one-season-out validation. The contribution is a bracket-centric modeling strategy that treats selection and seeding as consequences of one latent ordering process.

BUSINESS PROBLEM FRAMING

- We participated in the NCAA Men's Final Four Analytics Challenge competition focused on trying to use Predictive Analytics to determine optimal bracket seeding.
- NCAA tournament seeding is inconsistent and opaque; a data-driven ranking approach may improve accuracy.
- Seeding affects outcomes, revenue, and fairness for committees, teams, analysts, and fans.
- Challenges: Limited data, variable schedules, and incomplete seeding labels.
- Goal: To create transparent, consistent, fair, and replicable data-driven framework for decision support through combining classification and regression outputs into a unified prediction system

NCAA MEN'S FINAL FOUR ANALYTICS CHALLENGE

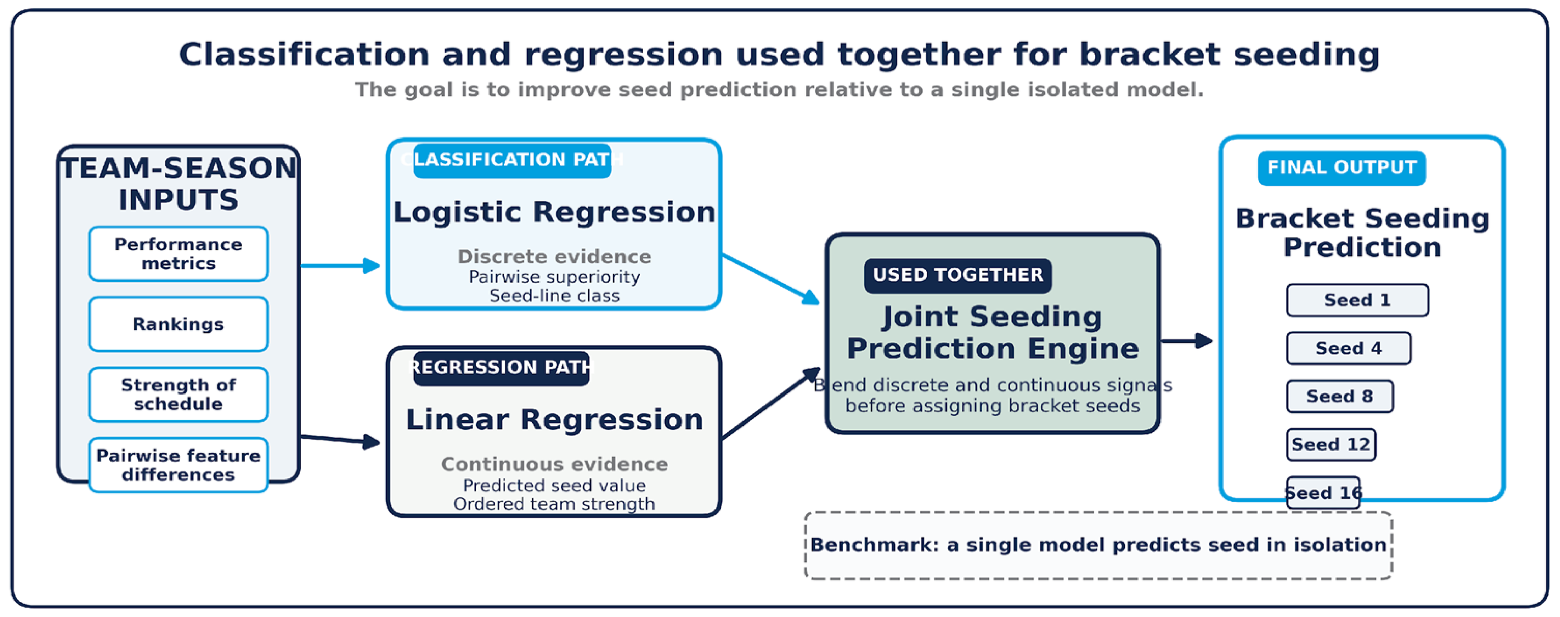




Research question: Can classification-type and regression-type models together improve the predictive seeding as compared to just performing one approach in isolation?

ANALYTICS PROBLEM FRAMING

- Framed as a learning-to-rank problem using pairwise classification and regression models to predict seeds.
- Assumes higher-seeded teams are stronger and that pairwise feature difference capture relative team quality.
- Uses Spearman and Kendall's tau for ranking accuracy and MAE for seed prediction accuracy.
- Ranking-first mirrors committee-style comparisons and improves interpretability by separating n from seed assignment.
- This approach reframes tournament seeding as a ranking problem through combining predictive models, and has an objective to minimize ranking error between predicted and actual NCAA seed.



DATA

NCAA team-season dataset with:
- Performance metrics
- Rankings
- Strength of schedule (SOS)

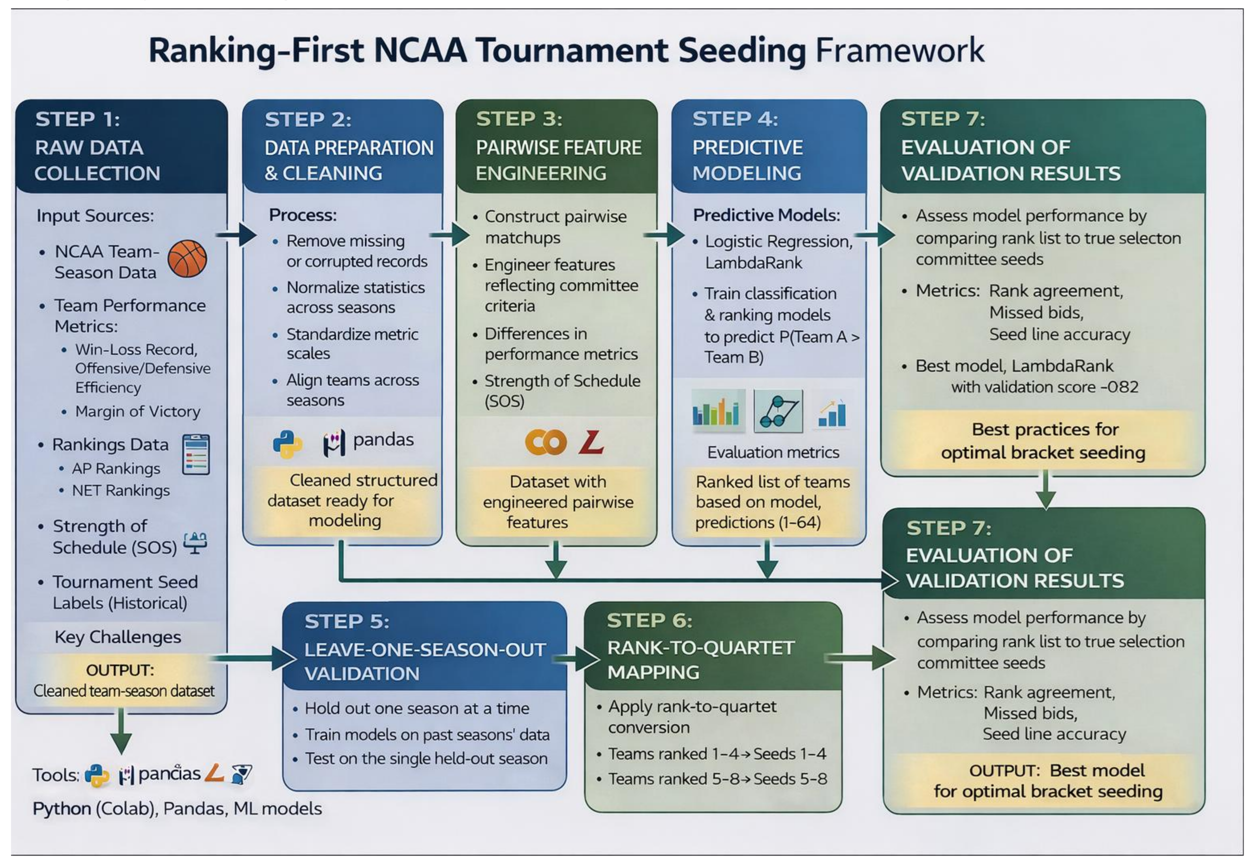
- Engineered **pairwise features** (team vs. team differences)
- Reflects **selection committee criteria**
- Supports **ranking → seed mapping pipeline**
- Teams with stronger records and higher-quality wins (Quadrant 1)

Overall Seed	The overall rank assigned for the 68 teams selected to compete in the NCAA
Bid Type	AQ = Automatic Qualifier. Assigned to team that won their Conference's tournament;
NET Rank	The NET Ranking (NCAA Evaluation Tool) is the primary metric used by the NCAA to
PrevNET	The NET Ranking (NCAA Evaluation Tool) is the primary metric used by the NCAA to
AvgOppNETRank	Average Opponent NET Rank. Ranks team's AvgOppNet compared to the rest of the
AvgOppNET	Average Opponent NET. Measures the average NET ranking of a team's opponents over
WL	Win-Loss; Number of Wins and Losses separated by a dash.
Conf.Record	Win-Loss against teams in the same conference
Non-ConferenceRecord	Win-Loss against teams in other conferences
RoadWL	Win-Loss with games hosted away from the teams home court.
NETSOS	NET Strength of Schedule. 1 = toughest, 364 = easiest
NETNonConfSOS	NET Strength of Schedule for Non-Conference games. 1 = toughest, 364 = easiest

Table 1. Data Dictionary

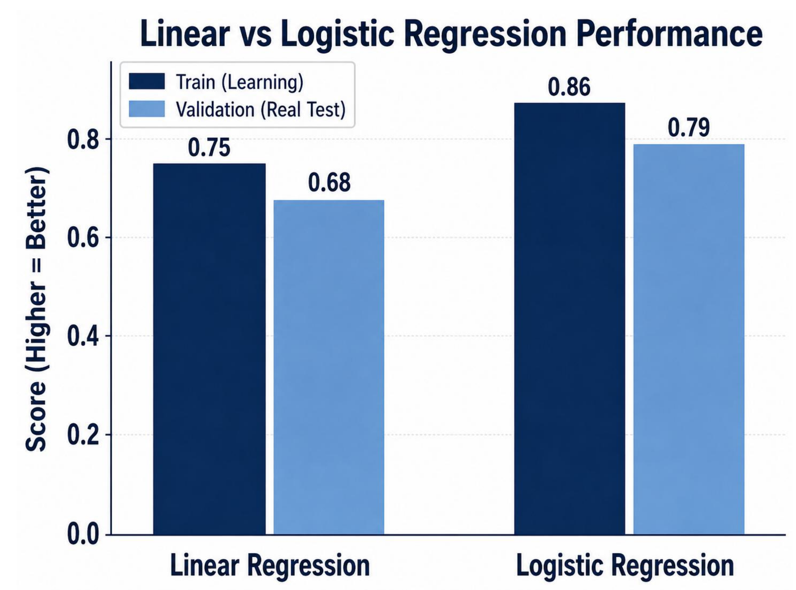
METHODOLOGY (End-to-End Predictive Pipeline)

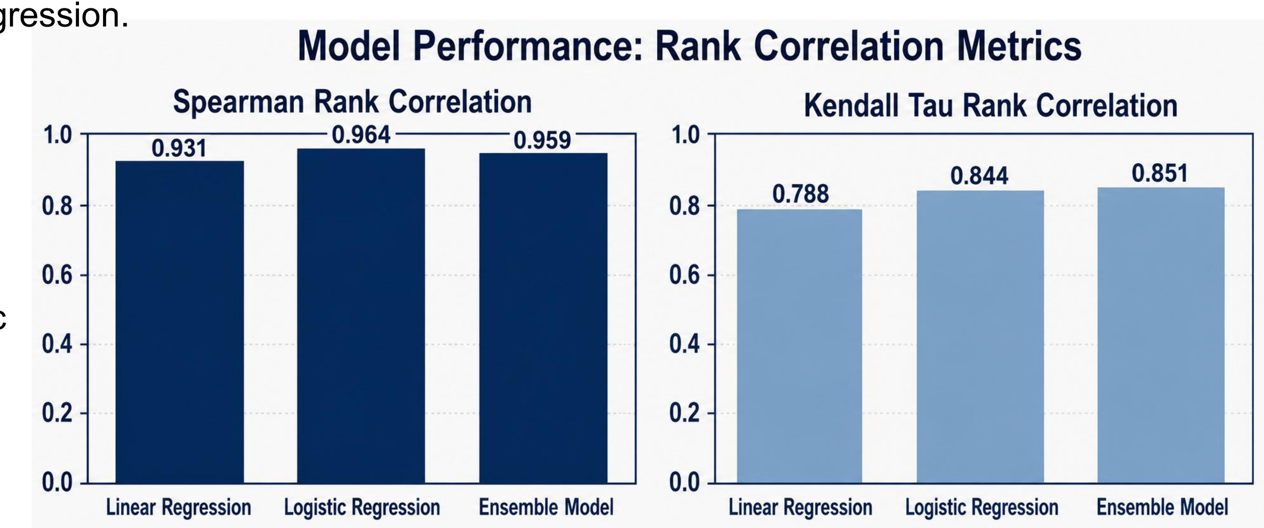
Data Prep: Cleaned data and created pairwise matchups
- **Modeling:** Predicted probability Team A > Team B → overall ranking
- **Validation:** Season-based (train past, test future)
- **Seeding:** Converted rankings into NCAA seeds (1–16)
- **Tools:** Python (Google Colab), Pandas, ML models



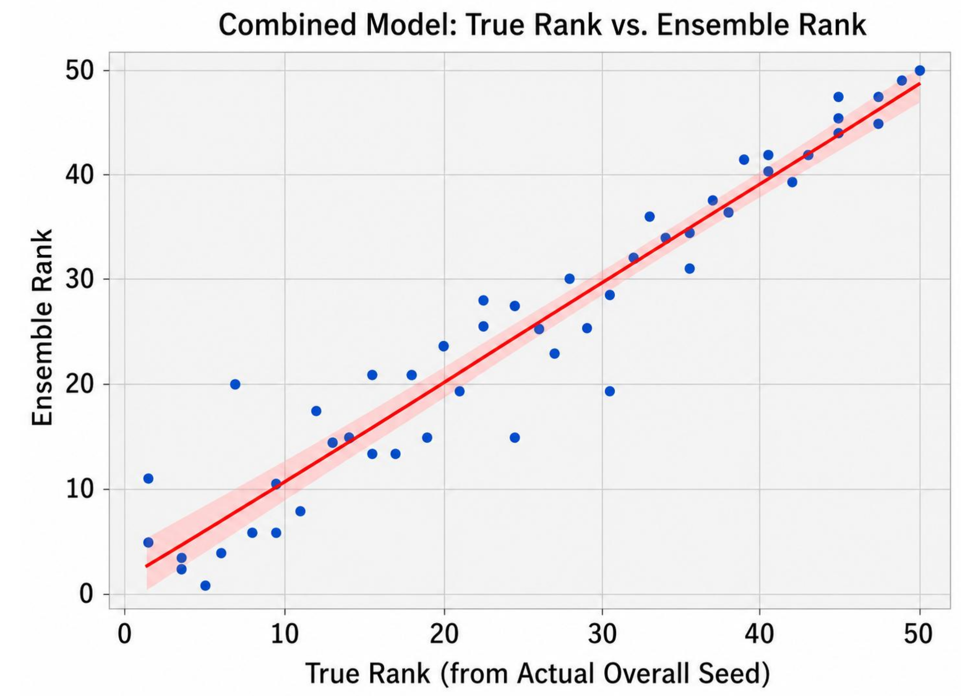
MODEL BUILDING & EXPERIMENTAL RESULTS

- **Linear Regression (Numeric-Based Model):** Uses quantitative performance metrics (e.g., NET ranking, wins/losses, strength of schedule) to predict team seeding with strong consistency across continuous variables.
- **Logistic Regression (Characteristic-Based Model):** Focuses on categorical and situational factors (e.g., quadrant wins, conference strength) to estimate probabilities of selection and relative ranking positions.
- The Logistic Regression model **(probability ranks) performed the best**, compared to the ranks generated by Linear Regression.
- Putting the ranks together **(Ensemble model) lead to a 0.70 increase in Kendall Tau Rank** performance compared to the Logistic Regression model.



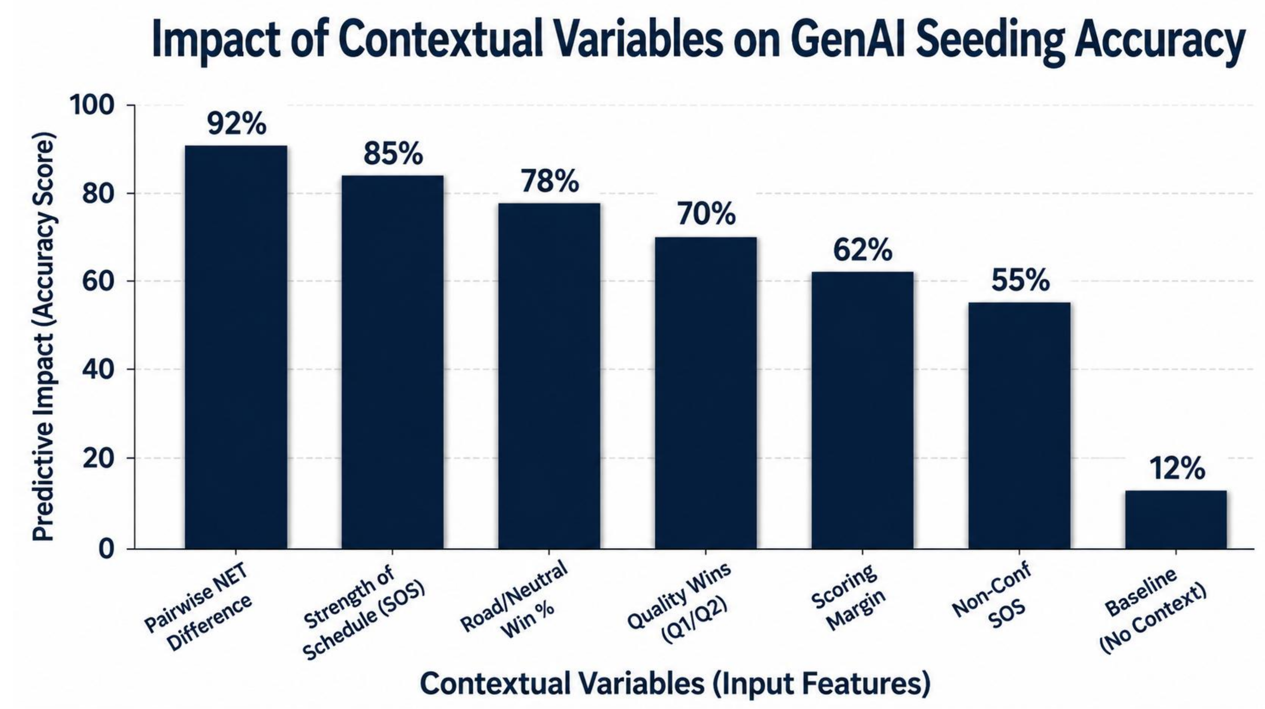


- **Model Comparison Insight:** Linear regression captures *magnitude and scale*, while logistic regression captures *likelihood and classification* — each provides a different lens on team performance.
- **Ensemble / Combined Model Advantage:** By integrating both models, the combined approach improves overall ranking accuracy, reduces bias, and creates a more balanced, data-driven seeding framework



DEPLOYMENT & LIFECYCLE MANAGEMENT

- This model can be used as a decision-support tool for the NCAA selection committee during tournament seeding and team selection.
- By combining multiple modeling approaches (linear and logistic regression), the system provides more stable rankings than any single model alone.
- The model helps reduce bias by offering a data-driven baseline for evaluating teams.
- The system would be updated annually using new season data to reflect changes in team performance and competition dynamics, and its approach is scalable and able to evolve over time



KEY TAKE-AWAYS

- **Ranking First Effectiveness**: We found that treating selection and seeding as a single latent ordering process provides a more cohesive framework than making isolated in-or-out predictions.
- **Feature Importance**: Engineered pairwise features – such as team vs team differences in performance metrics and strength of schedule were critical in capturing relative team quality.
- We placed 15th in the Final Four Analytics Challenge.
- Results show that combining predictive models improves reliability, accuracy, consistency, and frameworks for high-stake selection decision.
- Single-model approaches are less stable across datasets.
- Results demonstrated a decent alignment with actual NCAA selections.

Personal Development & Outcomes

DataCamp Certifications: Python fundamentals
Excel Associate Certification
SAS analytics badge