

TOUR DE FRANCE 2026

Predicting the next
Champion through
Machine Learning

CONNOR TOOMEY



PROBLEM

Predict the 2026 Winner

In thinking of this project, I wanted to see if I could predict the winner of the 2026 Tour de France after having followed the 2025 Tour very closely. Additionally, cycling is something that I greatly enjoy in my free time (when we aren't buried in snow).



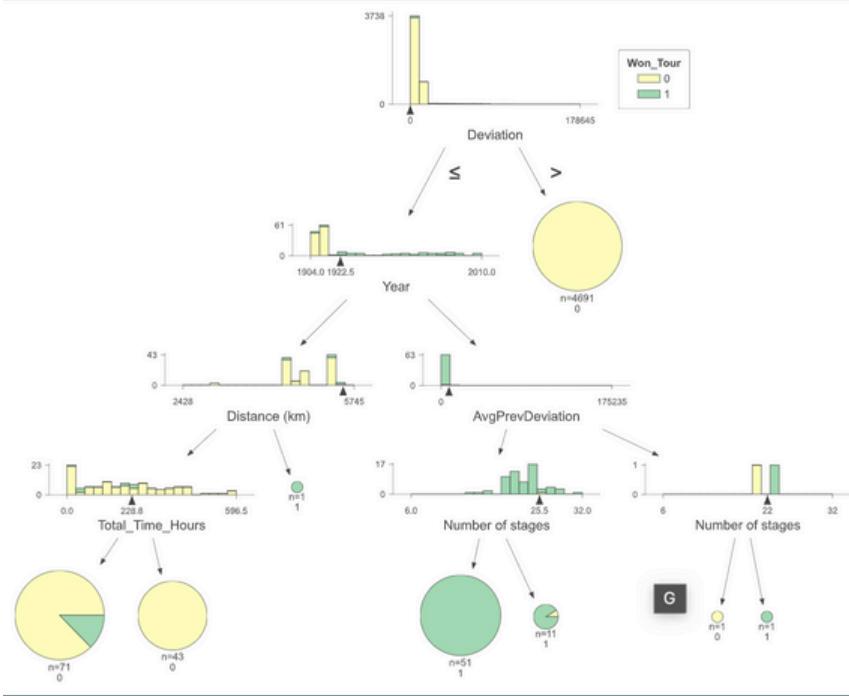
THOUGHT PROCESS

Different Disciplines

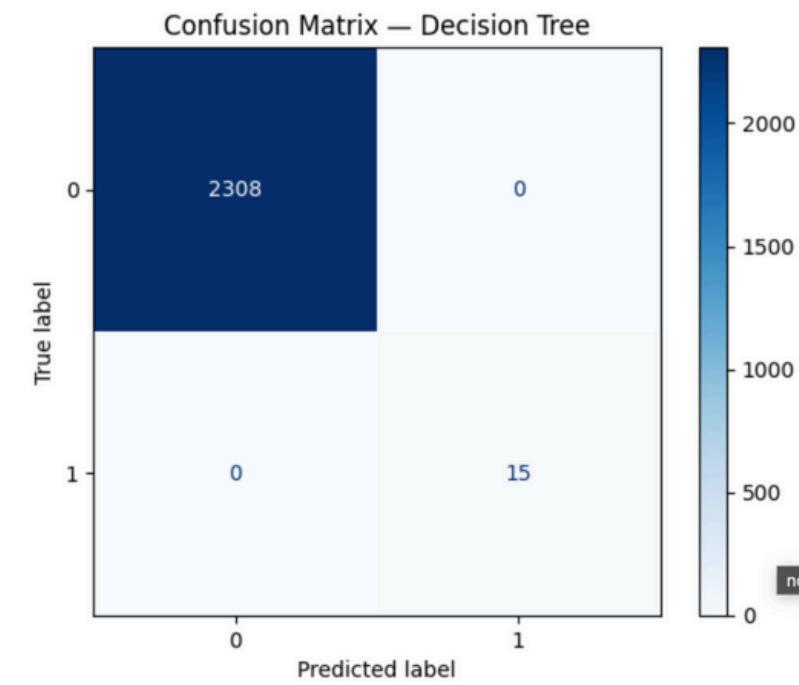


- Linear Regression
- Decision Trees
- XGBoost + Tuning
- Confusion Matrix
- Precision – Recall Curve

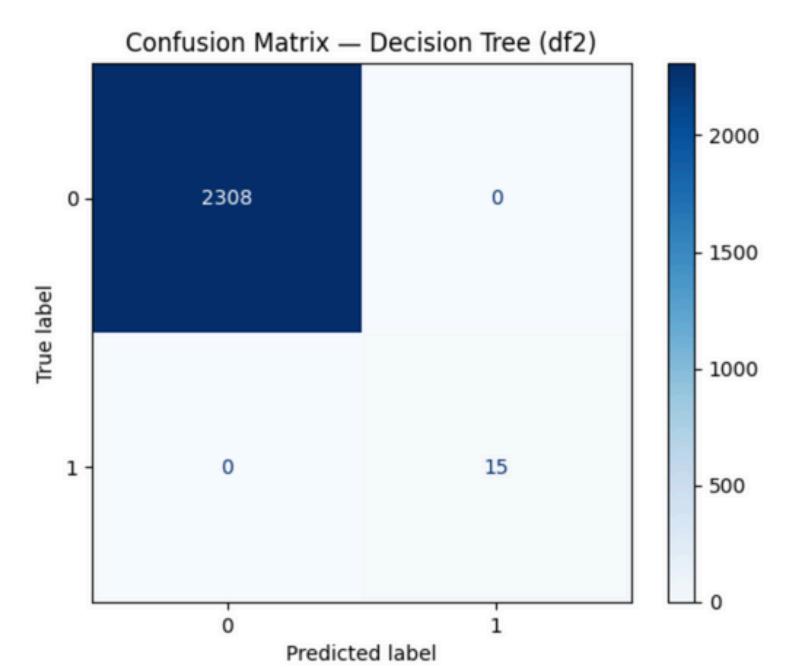
VISUALS - 1



Decision Tree for Predictors



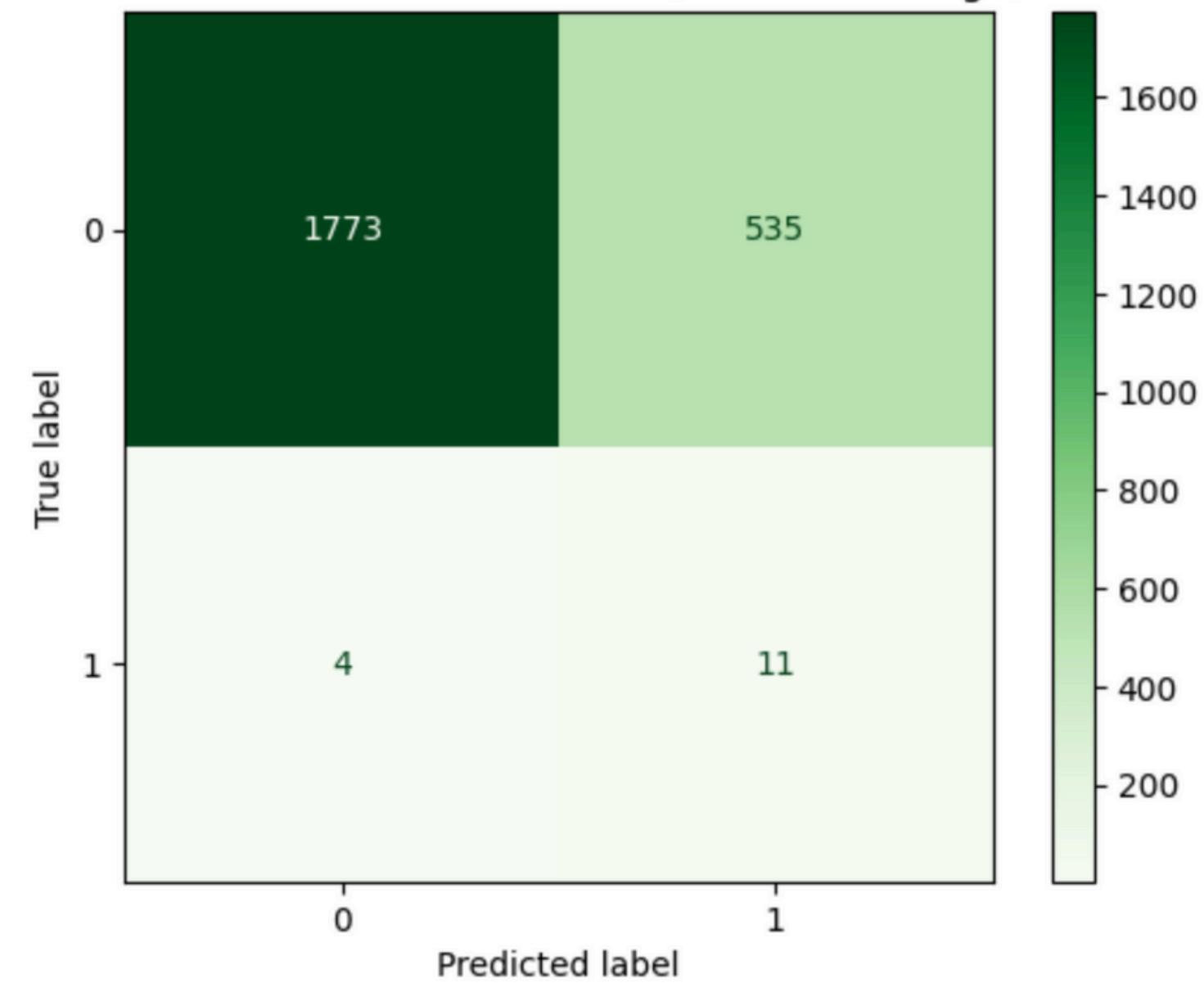
Confusion Matrix
based on Rider



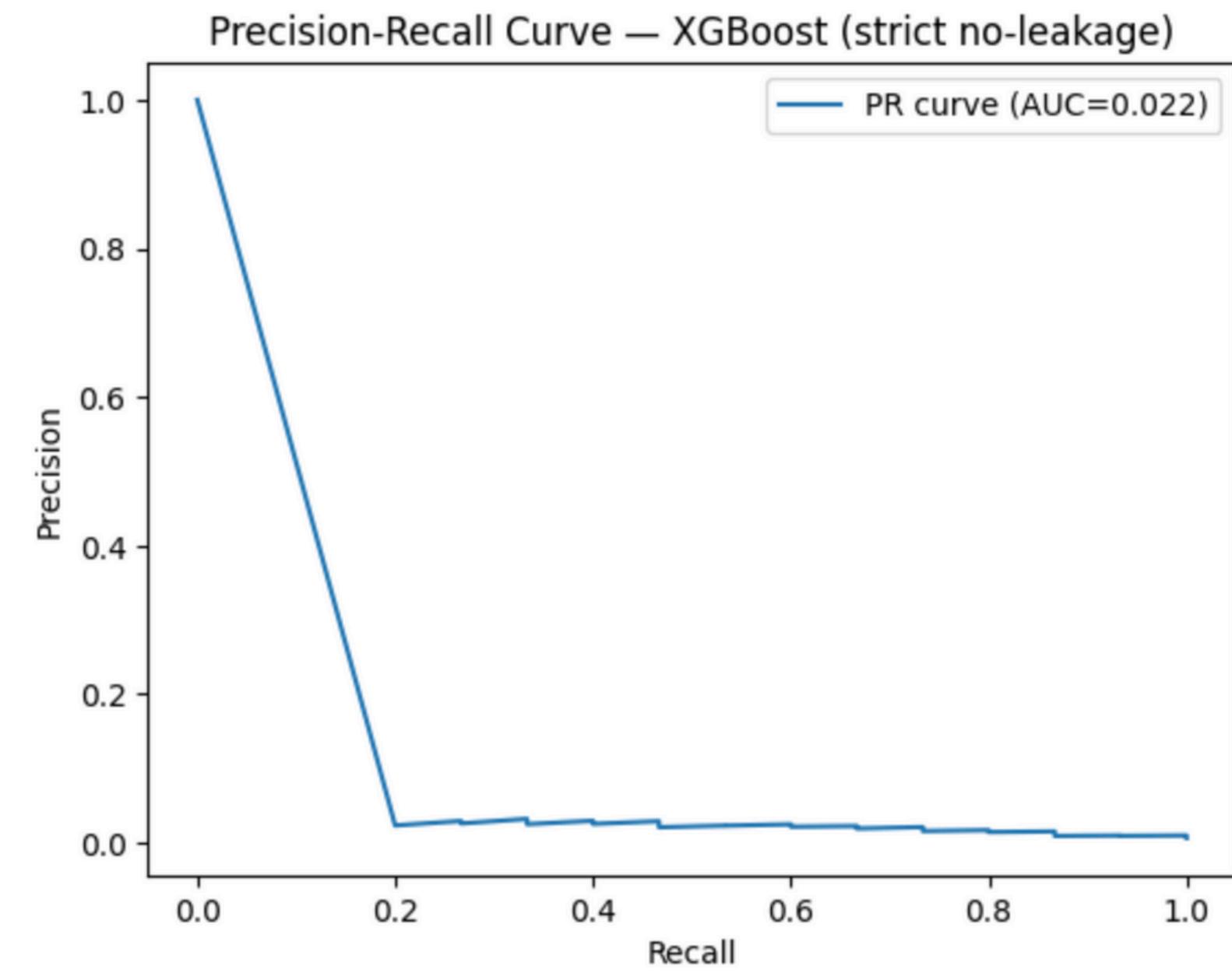
Confusion Matrix
based on Team

VISUALS - 2

Confusion Matrix — XGBoost (strict no-leakage)

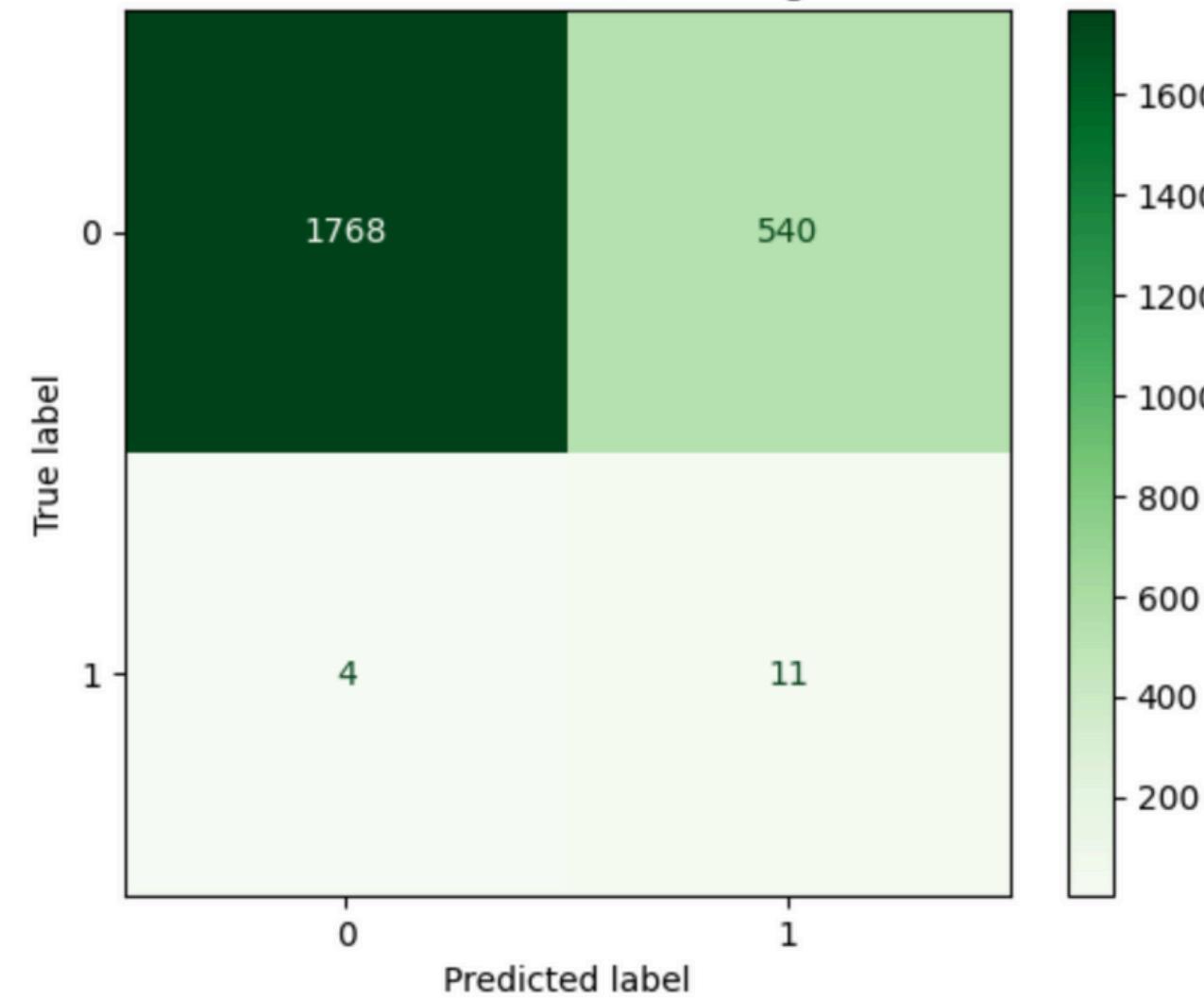


Precision-Recall Curve — XGBoost (strict no-leakage)

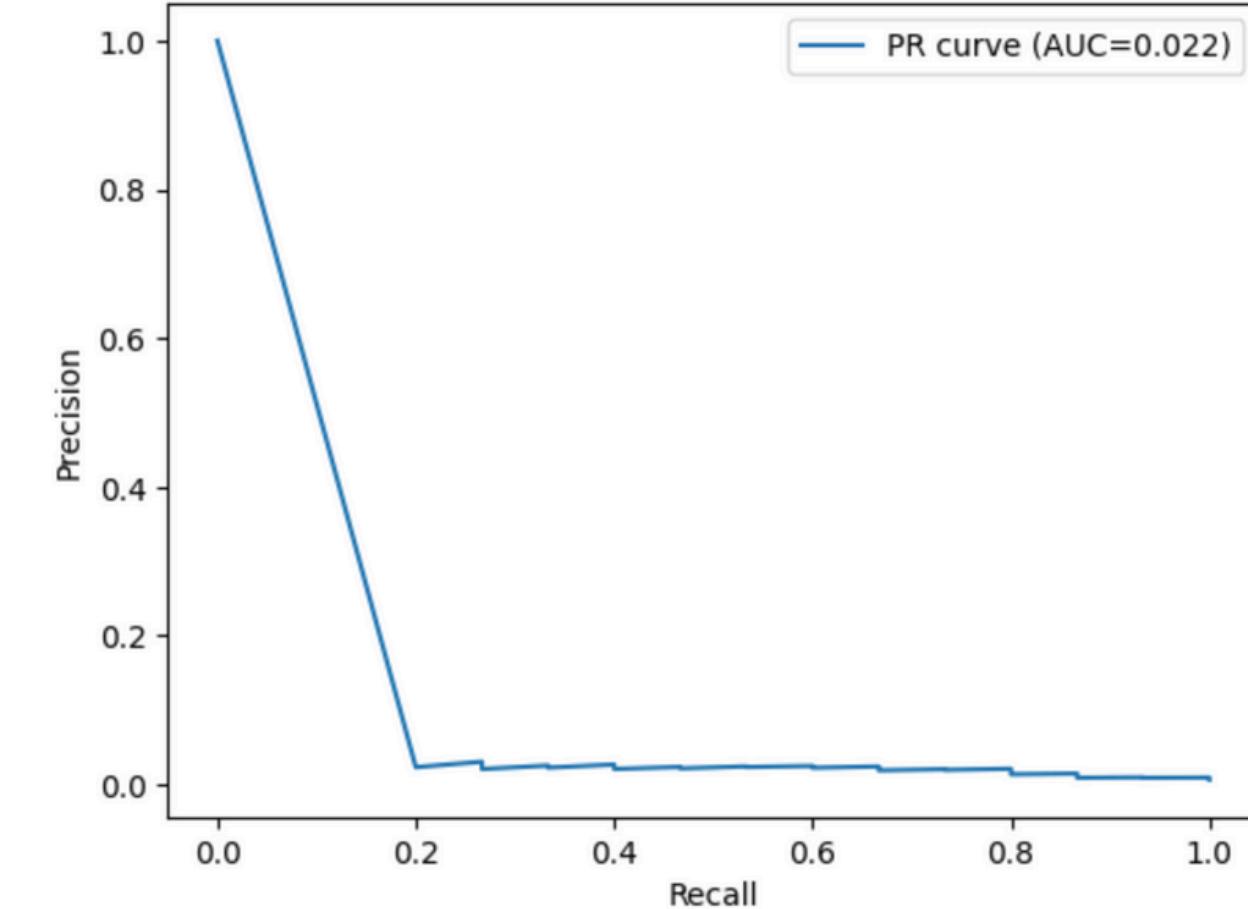


VISUALS - 3

Confusion Matrix — XGBoost (with stage-win features)

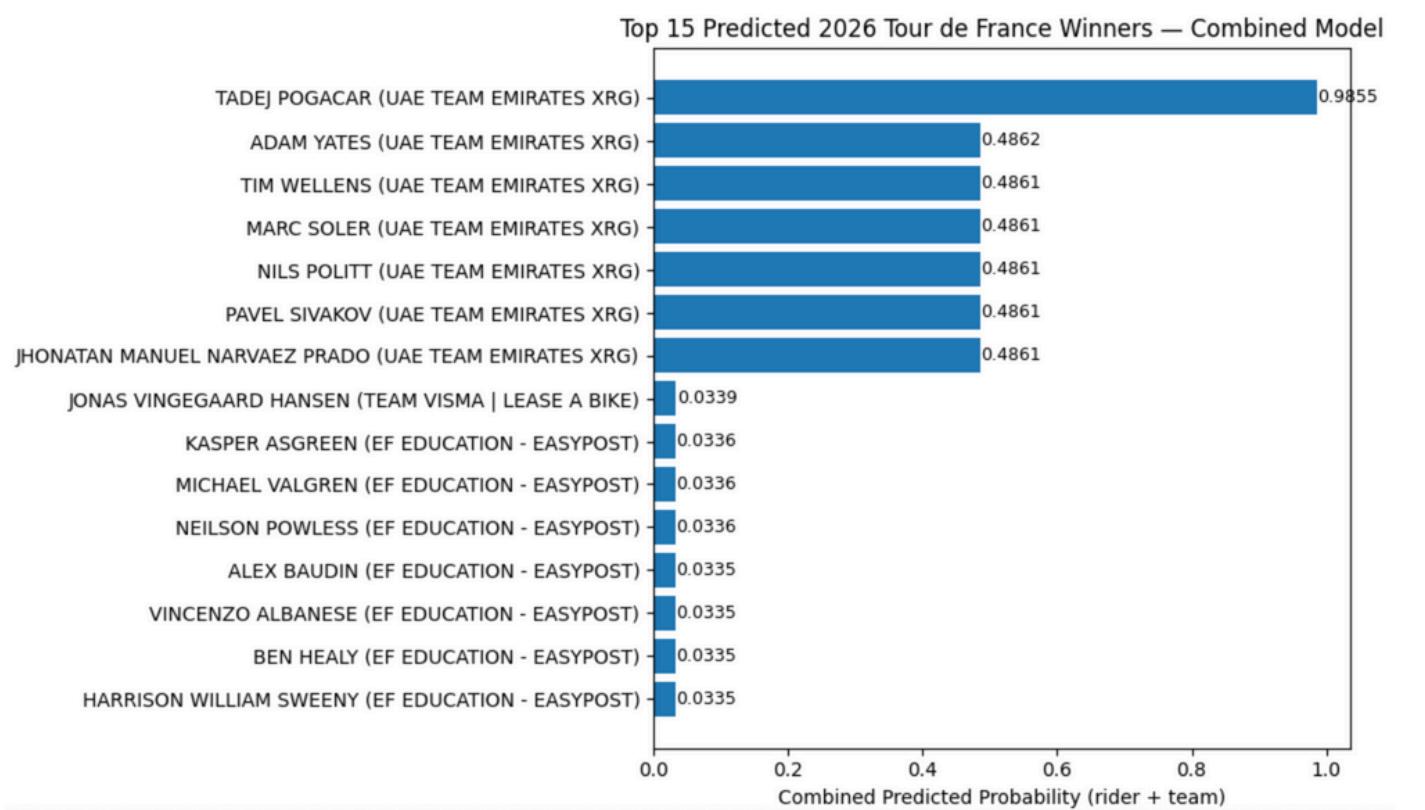


Precision-Recall Curve — XGBoost (with stage-win features)



YOUR 2026 WINNER

Tadej Pogacar



WHAT'S NEXT?

What more can we do?

If we were given more time during the Mod to really dive in and get into the thick of things, I would want to run a few different models using the same data. However, I would set parameters and features a little bit differently. I would work to predict the winner of the first stage of the tour, first three stages, and then if those worked I would try to predict stage by stage. This would be very difficult to do given the tracks/route the riders are on and the different specialities of the riders.



AU REVOIR ET MERCI

Connor Toomey

