

F1 Overtake Predictor

Katy Hosokawa



Motivation

Stakes: An overtake can have a 160 million championship difference

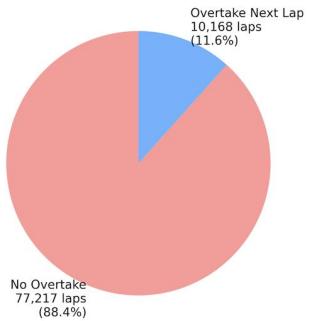
Problem: Team strategist have to make split-second data based decisions

Solution: A machine learning model that predicts next lap overtakes with **86% AUC-ROC**

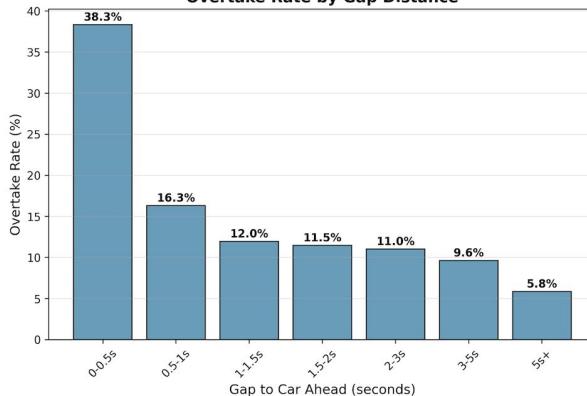


Data

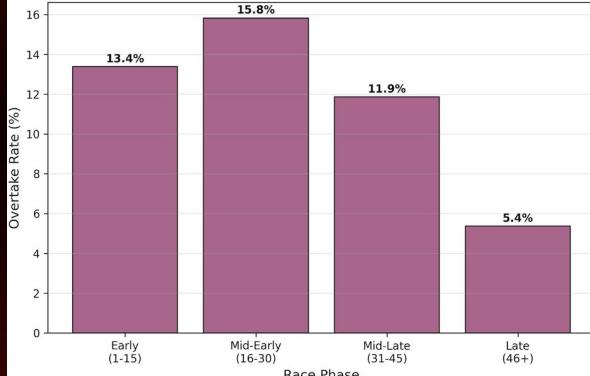
Dataset Scale: 87,385 F1 Race Laps
2021-2024 Seasons



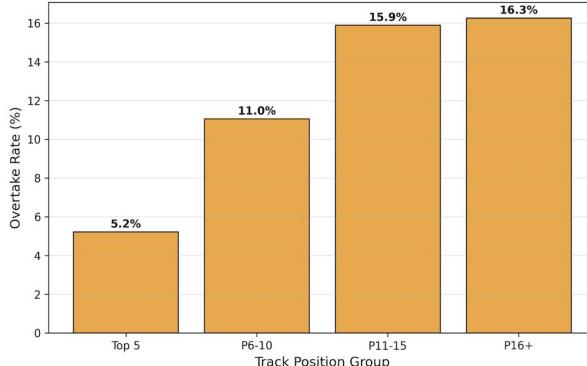
The Closer, The More Likely
Overtake Rate by Gap Distance



Racing Gets Aggressive Late
Overtake Rate by Race Phase



Midfield Battle is Real
Overtake Rate by Position



Preprocessing

Season 2021 - 2024 racing laps (88k records) with median/mode filled missing values, removed outliers, and rare category group encoding

Class Imbalance

11.7% positive class adjusted for with a 1:4 class weighting

The Model

Random Forest
500 Decision Trees

01

Matrix Factorization
Extracted 7 latent features
from **driver-track interaction matrix**

02

Hyperparameter Tuning
*Classification threshold,
class weighting, number of
SVD features and number
of trees*

03

Well-Rounded Features
*Permutively tried
combinations of recent
telemetry, track, driver,
weather data and
interaction matrices*

04

Features

Top 5/18 features ranked by importance from Random Forest

Gap To Car Ahead At Line (16%) *The distance to the next car at the end of the lap*

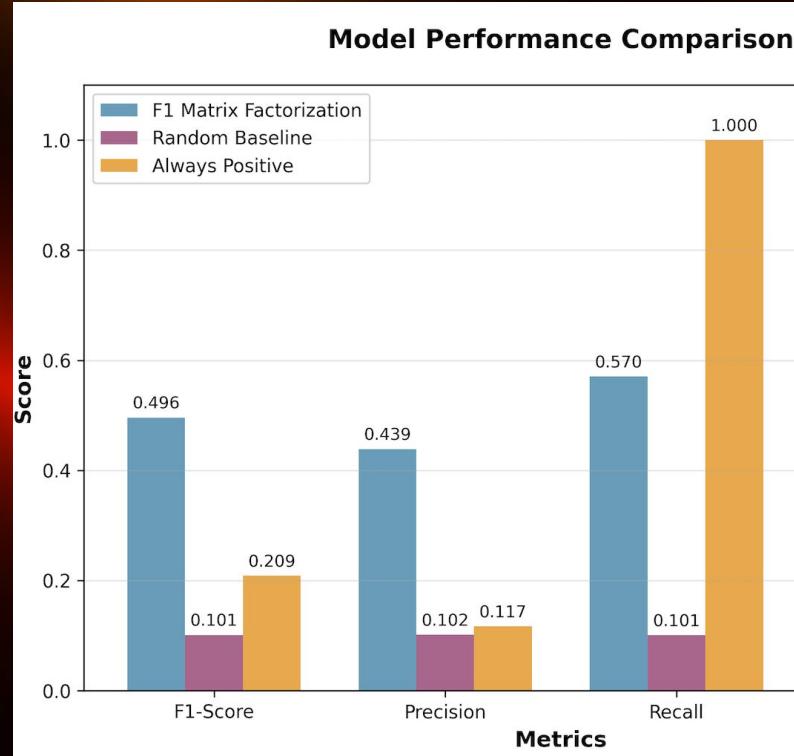
Lap Number (12%) *The current lap of the race*

Position (8.7%) *The current driver position / place*

Speed Average (7.6%) *Average speed of car throughout the lap*

Track Temp (7.4%) *Current temperature of the track*

Performance

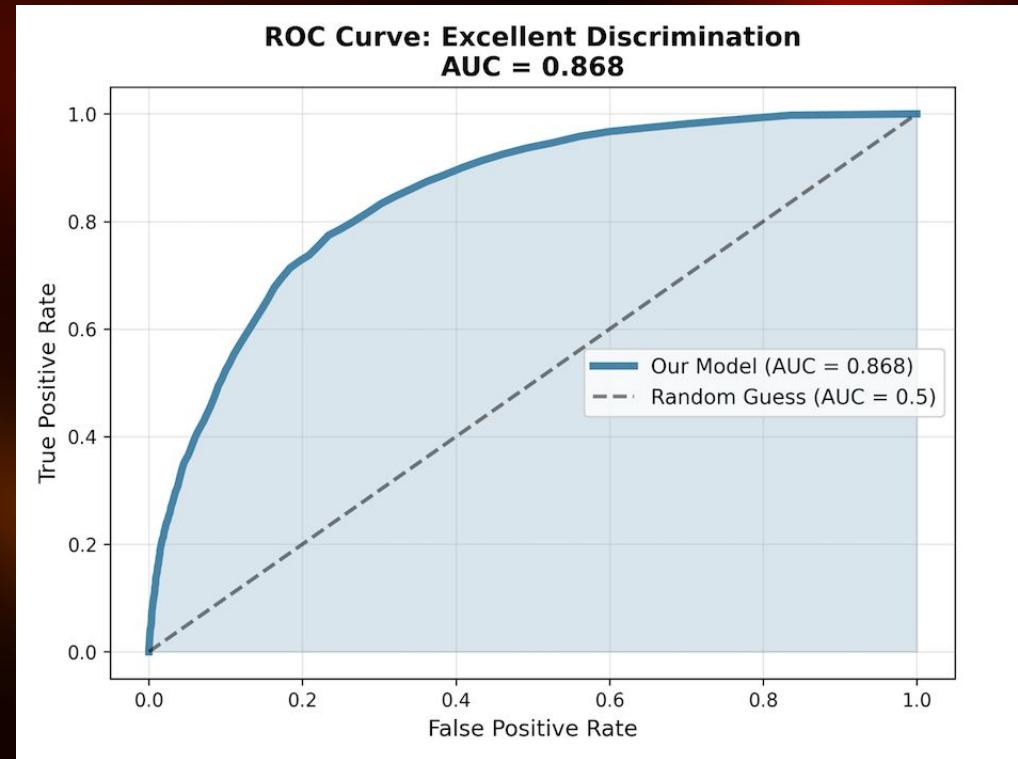


Evaluation

Train/test split with 3 K-fold cross-validation

Consistent F1 score and strong recall

AUC-ROC of 86.6% shows high distinguishing capability



Real-World Impact



*High speed decisions
that are backed by
data*



*Captures hidden patterns
not shown by explicit data
features*

*Interpretable decisions
that help explain
reasoning*

*Guide priorities for driver
improvement by
identifying key features*

Thanks!

