

Keywords - March 28, 2025

horse_id_valid

Ensures horse IDs are 7-digit validated strings.

class_num

Numeric representation of race class, crucial for race strength.

class_label

Describes class type (e.g., Handicap, Group) for race structure.

GoingStick

Going condition metric; useful when present.

presence_flags

Binary indicators for whether optional fields exist.

jockey_last_14_win_pct

Short-term form of jockey; now parsed from % string.

trainer_last_14_profit

Trainer's recent profitability; cleaned from string.

flatten_json

Core operation to convert nested JSON to flat DataFrame.

race_context

Metadata shared across all runners in a race.

form

Raw form string used to indicate past performance.

spotlight

Free-text insights about runner; potential NLP source.

headgear

Indicates presence of blinkers, cheekpieces, etc.

extract_features_v2

Main function converting raw JSON into model features.

inject_country_and_flatten

Injects country info and flattens all runners in JSON.

percent_parser

Utility to convert '13%' into usable float values.

string_id_cast

Forces IDs to be treated as strings instead of floats.

field_overwrite_bug

Previously caused class_num/label to vanish; now fixed.

data_auditing

Manual checks for presence, structure, and correctness.

zscore_normalization

Creates standardized versions of ratings within a race.

race_class_parser

Regex-based parser to extract numeric class and labels.