

Daily Dev Summary - Transformer Racing Model

Date: 2025-04-02 18:58

Quick Summary of What You've Built So Far

- A transformer-based architecture tailored to racing data.
- A full preprocessing and inference pipeline - clean, modular, and testable.
- Integrated NLP signal (comment + spotlight).
- Dynamic race batching with masking for variable-length races.
- A flexible training loop with checkpointing, logging, and loss tracking.
- Evaluation tools with winner prediction, ranking comparison, and plots.
- The ability to extend this from batch experiments to full-scale training.
- A project worth showing off and building on.

What's Next (Ideas)

- Expand training data month-by-month (as new races/results come online).
- Add Top-K evaluation metrics (e.g. Top-1, Top-3 accuracy).
- Begin testing on live or upcoming races.
- Try alternate target strategies (e.g., position regression, multi-label win/place/show).
- Evaluate model ensembling or alternate model architectures (MLPs, CNNs, tabular transformers).
- Auto-generate:
 - * Documentation
 - * Evaluation dashboards
 - * GitHub release workflows
 - * Deployment (TorchScript, ONNX, Streamlit app, etc.)

Data Matching Limitation

Current inference vs. results merge achieves ~66% match. For improvement:

- Work on cleaning course and name mappings (handle misspellings, aliases).

- Consider expanding scrape coverage across 1-day before and after.
- Consider removing countries with persistent result lookup issues.
- Eventually, focus on UK/IRE if needed for accuracy and maintainability.