

Horse Race Predictor MVP - Implementation Checklist

Core Features

Backend API

- ✓ Integrate Oracle Engine's prediction API pattern for model loading and inference
- ✓ Create Python wrapper for rpscraper to fetch live racecards (today/tomorrow)
- ✓ Create tRPC procedure to fetch available races
- ✓ Create tRPC procedure to fetch horses in a selected race
- ✓ Load trained LightGBM model (`ranking_model.pkl`) into server memory
- ✓ Create tRPC procedure to generate predictions (top 4 horses in order)
- ✓ Preprocess race data to match model input format (442 features)
- ✓ Return predictions with confidence scores
- ✓ Integrate data models from Android app for type safety

Frontend UI

- ✓ Create race selection interface (meet/track dropdown + race selector)
- ✓ Create prediction display showing 1st, 2nd, 3rd, 4th place horses
- ✓ Add loading states during prediction generation
- ✓ Display model confidence/probability for each prediction
- ✓ Add error handling for API failures

Data Integration

- ✓ Research and identify live data sources (rpscraper, Betfair API, or TAB NZ)
- ✓ Create data ingestion service to fetch live racecards
- ✓ Implement scheduled job to refresh race data every 30 minutes
- ✓ Set up rpscraper Python wrapper in Node.js backend (child_process)
- ✓ Cache racecards to avoid repeated scraping

- ☒ ~~Validate scraped data before passing to model~~
- ☒ ~~Integrate feature engineering from Oracle Engine~~
- ☒ ~~Map racebase scraper output to model feature requirements~~

Deployment

- ☒ ~~Test MVP with live data~~
- ☒ ~~Expose public URL for user testing~~
- ☐ Document API endpoints and usage

Prediction Accuracy Tracking (NEW)

- ☐ Design prediction tracking schema
- ☐ Create database tables for predictions and results
- ☐ Build race results fetcher service
- ☐ Implement accuracy metrics calculation
- ☐ Create backend API endpoints for tracking
- ☐ Build performance analytics dashboard
- ☐ Integrate dashboard into MVP
- ☐ Test end-to-end tracking and analytics

Completed Tasks

- ☒ ~~Initialized web-db-user project~~
- ☒ ~~Trained LightGBM Ranker model (NDCG@3: 0.8579)~~
- ☒ ~~Consolidated 396K+ horse racing records~~