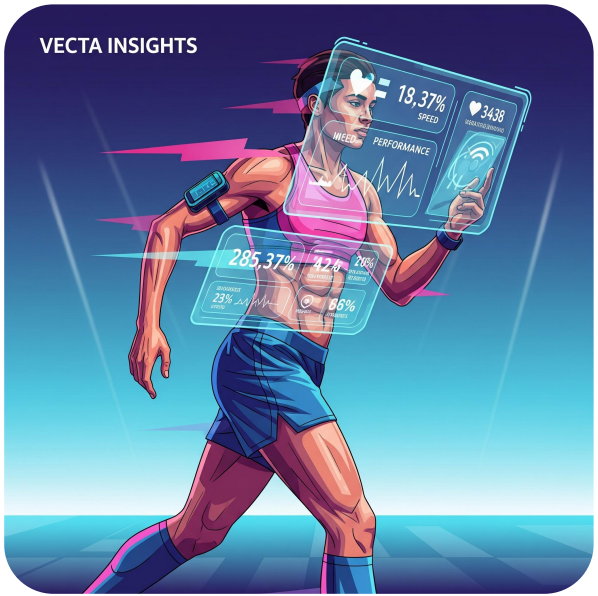
HyroxGPT: Personalized Feedback Generator for Hyrox Participants

HYROX

Combines both running & functional workout stations, where participants run 1km, followed by 1 functional workout station, repeated eight times

HyroxGPT

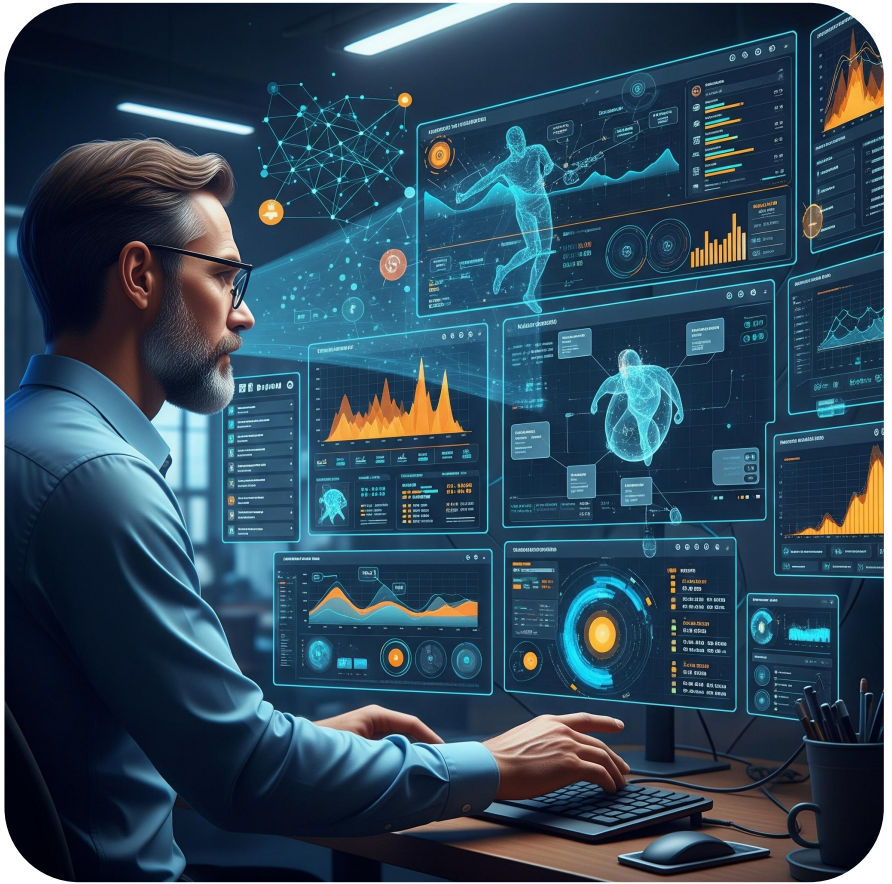


Aim to provide personalized feedback to Hyrox athletes considering the next questions:

- Did the athlete performed better or worse than expected overall?
- What type of athlete is the person grouped with (e.g. age, gender).
- In which exercises did the athlete underperformed vs their peers?

How to solve it?

- Using Machine Learning to explore participant datasets.
- Analysing metrics compared to peers via clustering.
- Predicting the expected time to finish the race and compare it with real result.
- Fine-tuning LLM models to be able to generate a human feedback to users based on their metrics.



🧠 Conclusions:

Throughout this project, I explored different approaches to generate personalized performance feedback for Hyrox athletes.

- I’m very satisfied with the machine learning part. Training models like Linear Regression and LightGBM gave solid and interpretable results based on real race metrics.
- The first fine-tuning attempt with GPT-2 was surprisingly smooth. It was easy to train, but the generated feedback still lacked dynamism and human tone, making it less user-friendly than I would like for a real solution.
- Moving on to fine-tuning Mistral-7B using LoRA, I ran into more realistic development challenges.
    - The training process took over 2 hours, and I used up my entire GPU quota on Google Colab.
    - After that long wait, I realized I had forgotten to configure the offload_folder parameter and as a result, I couldn’t continue the generation phase since the model couldn’t load properly, and I was blocked from making further progress without GPU.
- Considering my dataset only had ~5,000 rows, scaling this to a full product with historical Hyrox results would be a huge undertaking.
- I would have loved to finish training the model fully with Mistral, as I believe it could deliver much richer, human-like feedback with the right resources.

Next Steps? 🤷‍♀️

1. Complete the Mistral Fine-Tuning
    a. Train again and finish it
    b. Evaluate generated feedback and compare with GPT-2 baseline.
    c. Optimize prompts and test on new athlete / participant profiles.
2. Explore a RAG-Based System (Retrieval-Augmented Generation), A scalable alternative to full fine-tuning for large, evolving datasets.
    a. Build a vector store from scraped historical Hyrox results.
    b. Vectorize context: athlete metrics + cluster profile.
    c. Let the LLM generate feedback based on retrieved nearest athletes:
        i. *Athletes in your group are 30s faster in burpees. Here's how to improve…*