PPO Market Bot

Mathys Vinatier

Seoul National University

October 14, 2025

Data Tick Update

	Signal Capture	Noise Level	Transaction Cost	Computation	Recommended
1-second					not
1-minute					2nd choice
5-minutes					1st choice
15-minutes					not

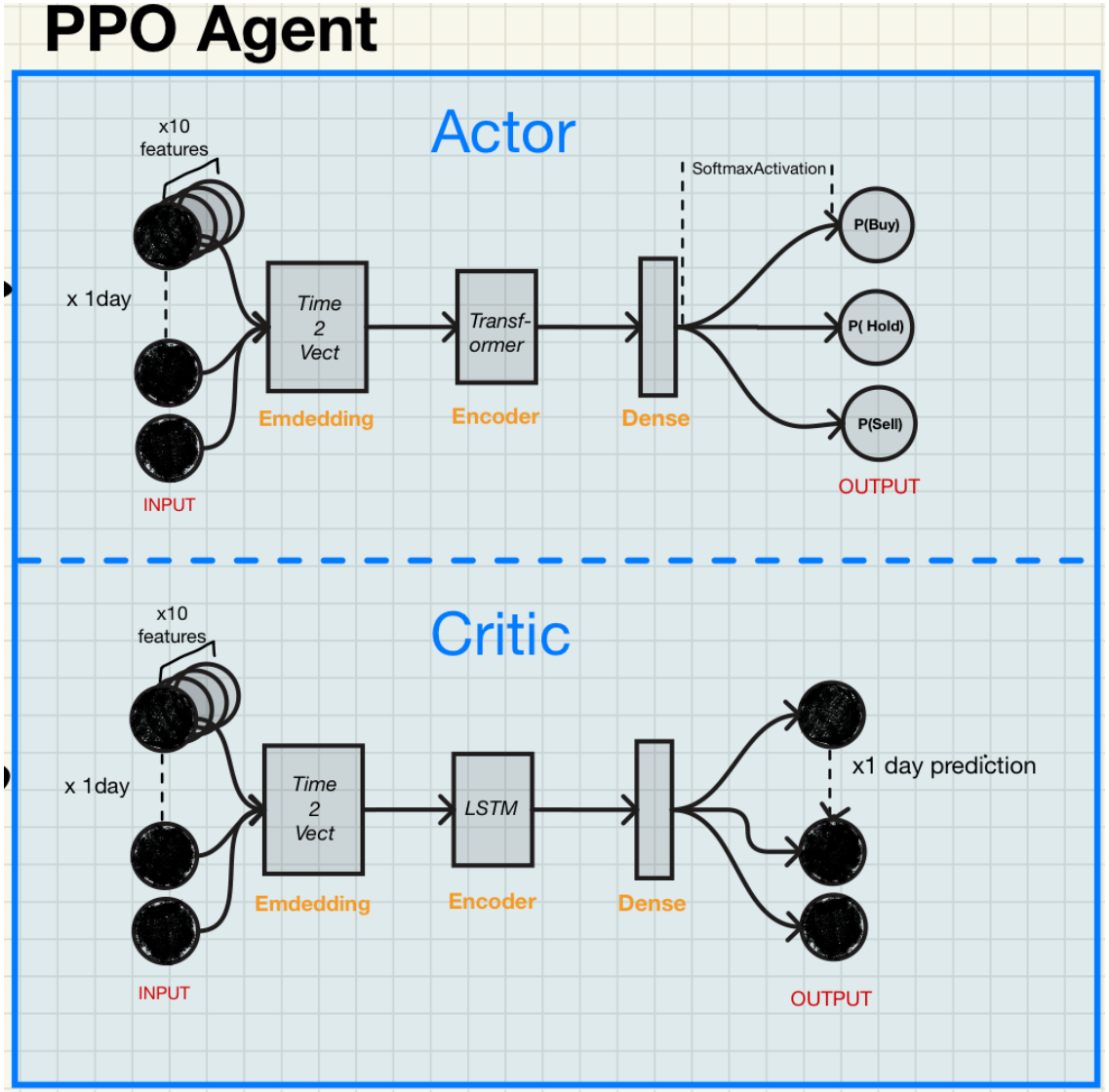

Figure 1: PPO agent architecture for the market environment

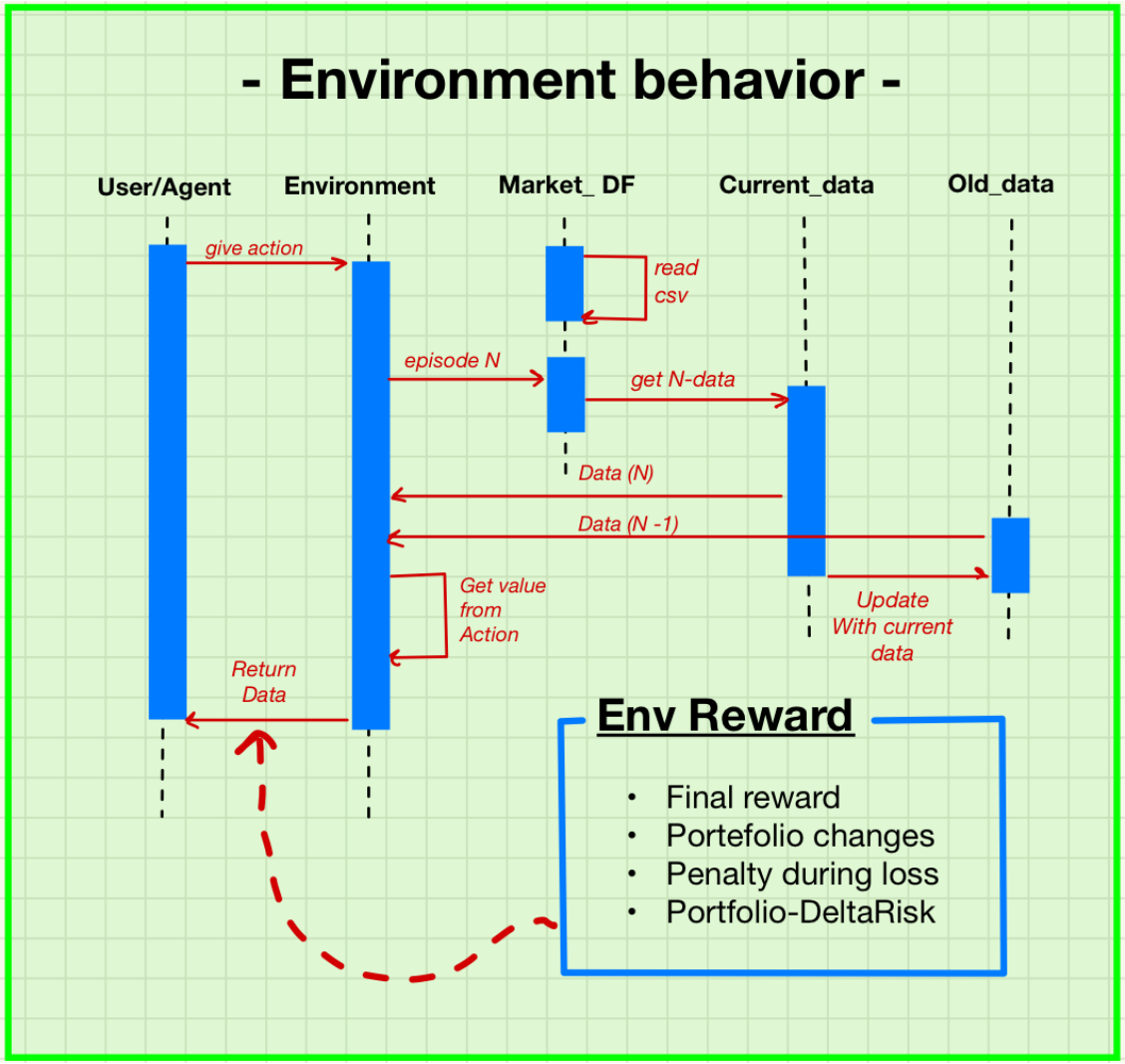

Figure 2: Sequence diagram of the environment's behavior

Fine tuning the reward

- **Final Reward :** total profit over a given period

- **Portfolio Change :** difference between the portfolio value before and after an action is taken

- **Penalty Loss :** negative reward whenever the agent make a losing trade

- **Portfolio Delta Risk :** risk-reward based on the volatility associated to an action

Suggestions and Future Work

Next steps :

- Add **S&P500** features into the environment (5-minutes tick data)
- Refine the **Actor** deep learning architecture
- Develop the **Critic** deep learning architecture
- Implement **performance analysis** tools
- Use the performance metrics to **fine-tune** the **reward**
- Add **risk management** features

Trade Station API Service

API futur updates :

- Reduce the update interval from 15 minutes to a **5-minute basis**
- Add a dedicated database for **range-based data requests**
- **Implement functionality** to retrieve and transmit a range of data (both server- and client-side)

Thank You