Apr 17 · 3 min read

Agents predict the future: unveiling the Olas Predict Leaderboard on HuggingFace

How good is AI at predicting the future?

Using LLMs to predict future events has been largely underexplored. Today, we're excited to introduce the Olas Predict Leaderboard on HuggingFace, a dynamic dashboard designed to explore just that! It showcases that AI prediction tools - Python scripts implementing various prediction strategies leveraging AI models – are predicting the future with up to 75% accuracy. These tools are actively used by autonomous agents participating in the perhaps largest economy of AI agents, Olas Predict, where they are trading real value in over 450,000 transactions so far! Created in a collaboration between Valory and Naptha, as a contribution to Olas Predict, this benchmark evaluates the predictive abilities of these AI tools.

While the HuggingFace and Olas communities differ in specialisms, they share many similar underlying values. Both are strong proponents of open source, and create tools to put the power of AI in the hands of the community. In this blog post, we explore how Olas can unlock the expertise of the HuggingFace community.

Why a HuggingFace Leaderboard?

Benchmarks are crucial in machine learning and data science, providing a standardized framework to evaluate and compare various models and algorithms (e.g. The Open LLM Leaderboard). They enable researchers and developers to measure progress, identify strengths and weaknesses, and drive innovation by setting clear goals within the community. The Olas Predict Leaderboard on HuggingFace exemplifies this by standardizing the evaluation of prediction tools, thereby enhancing our understanding of AI-driven predictions. HuggingFace, a leading benchmarking platform, simplifies the process for enthusiasts and developers to test and refine their tools, advancing our ability to predict the future, particularly, in the context of the Olas prediction agent economy.

Overview of the Dashboard

The dashboard features a leaderboard that ranks AI-powered prediction tools based on accuracy, cost efficiency, and other vital metrics. It not only highlights the performance of each tool but also gives users the ability to run these tools against the benchmark dataset directly from within the interface of the HuggingFace space. The tools employ different strategies and models. Strategies range from RAG to reasoning to chain-of-thought, which browse, scrape and aggregate data among many other functions. Models employed include the most advanced closed-source LLMs from OpenAI and Anthropic, as well as open-source alternatives from teams including Nous Research and Databricks. The backbone of this benchmark is the refined Autocast dataset. This dataset contains questions of different categories. The dataset has been adapted specifically for evaluating the performance of Olas's predictive tools which focus on binary true/false questions.

Olas Predict

These tools are not merely experimental; they are actively used by autonomous agents to trade real value on prediction markets. These "Trader" agents use AI tools accessed via Mechs—permissionless marketplaces designed specifically for agent use cases—to determine the probabilities of events and

make informed trading decisions. Mechs offer a wide range of AI-powered prediction tools that all agents can use. Every tool featured in the benchmarks is available through these mechs.

Participate in Shaping the Future

The top tool on the leaderboard currently boasts an impressive 75% accuracy rate. Think you can do better? Create your own prediction tool, run the benchmark, and improve upon the current leader. Visit the Olas Mech Build Path to get started. You can then get your tool added to mech services to be used by prediction agents, and qualify for potential rewards. This is your chance to contribute to a cutting-edge field and see your work in action.

Ready to see what the future holds? Visit the Olas Predict Benchmark Dashboard on HuggingFace now and start making your mark!

About Valory

Valory is a research and deployment company at the intersection of crypto and AI. Specifically, the premier creator of open-source frameworks for co-owned AI.
https://www.valory.xyz/

About Olas

Olas is the network for co-owning AI. Olas enables everyone to own a share of AI, specifically autonomous agents. One of the first Crypto x AI projects, founded in 2021, Olas offers the composable Olas Stack for developing autonomous AI agents, and the Olas Protocol for incentivizing their creation and co-ownership. Olas' mission is to incentivize and coordinate different parties to launch autonomous agents that form entire AI economies serving all humans.
https://olas.network/

About Naptha

Naptha is a platform that makes it easy to deploy and operate AI workflows and multi-agent systems across a network of nodes, to solve real-world tasks contributed by users.
https://www.naptha.ai/