

Introduction

Playing the stock market is one of the many frontiers of applied artificial intelligence. The problem of optimizing asset allocation within a portfolio to yield a return above the market is a hard problem. Current approaches include multi-agent systems in which the task of gathering data, predicting trend and choosing assets are divided between specialized cooperative agents that simulate the environment of an investment fund.

This paper builds upon existing models to create ORACLIA, a multi-agent system for portfolio optimization in the Brazilian Stock Market (Bovespa). A multi-agent architecture was defined and implemented with the use of machine learning classification models to generate buy and sell signals based on the success or failure probability of a predefined strategy.

ORACLIA: Model and Design

The usual approach when dealing with stock price trend prediction is using non-linear regression models to predict future stock prices. Regression is a complex problem, with unsatisfactory results when dealing with stock prices. Predicting the price three days ahead (or three minutes ahead with intraday stock price) is a much harder problem than predicting one day ahead.

To avoid this pitfall, ORACLIA turns predicting stock price trend into a binary classification problem. In order to do that, we first predefine a strategy, and an asset. The simplest strategy, called swing trade, is buying an asset, and selling it if the price increases or decreases $x\%$ within the next n days; x/n will define the expected profit per day. Riskier strategies will be successful less often, making it harder to predict.

With a strategy defined, the agent evaluates the entire training set, and for each observation creates a binary label, which indicates whether the strategy would have been successful or not if applied on that day. The success or failure of the strategy will be the label learned by the classifier. The goal of the classification problem is, given the variables for the current day, define whether applying a strategy on a specific condition will be successful or not.

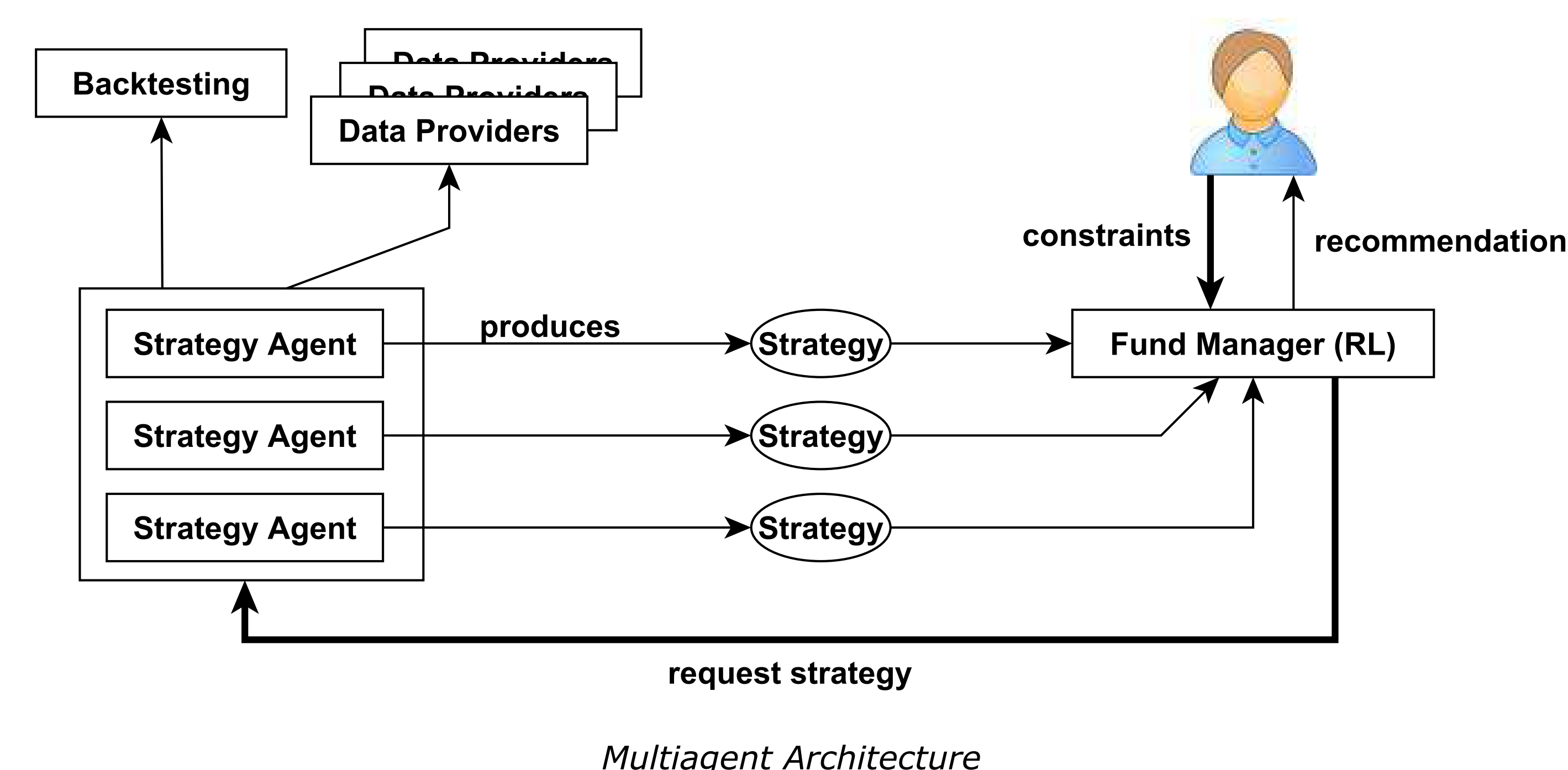
The predictions are used to make decisions on which asset to buy or sell and when. There is an inherent trade-off as we are disregarding additional information that would be obtained in the regression-based approach, such as the expected price of the asset, to obtain a simpler model which can be trained to higher precision.

The multi-agent system is composed of three layers:

Data Provider: Gather and transforms data from multiple sources

Strategy Agent: Turns data into predictions of the success or failure of a pre-defined trading strategy

Fund Manager: Uses predictions generated by strategy agents to make asset allocation decisions in order to optimize a portfolio's return on investment.



Experiments

A prototype of ORACLIA was tested with the Brazilian stock market (Bovespa). Data from the five stocks with higher weight in the Ibovespa index were used in the experiments: ABEV3, BBDC4, ITUB4, PETR4, and VALE5. The training data considered the period from 2002 to 2014, while the evaluation data considered the period from Jan. 2015 to Oct. 2016.

Features			
Trade Info	Fundamental		Technical
Open	Selic	Bovespa total volume	Moving Average <10, 20, ..., 60> days
Close	Exchange Rate USD	Bovespa index	Bollinger Bands <10, 20, ..., 60> days
High	BM&F Gold gramme	Nasdaq index	
Low	International reserves	Dow Jones index	
Volume	Foreign exchange operations balance		

Features collected by data providers and used at classification task

For each stock, a strategy agent was implemented, using a swing trade strategy and a k-nearest-neighbors classifier. In backtesting, the strategy agents achieved precision as high as 80% in predicting whether or not a strategy would be successful.

The fund manager agent simulated an investment portfolio in the period of 2015-2016, using the insights provided by strategy agents. The maximal return of investment achieved in the period is eight times higher than the IBOVESPA index gains or any of the traded stocks treated separately.

Classifier Score: Precision*					
Asset	ABEV3	BBDC4	ITUB4	PETR4	VALE5
Mean	0.87	0.73	0.74	0.86	0.80
Standard Deviation	0.07	0.11	0.08	0.05	0.06

* Cross-validation results (10 folds, stratified)

Performance of a kNN classifier, after extracting principal components.

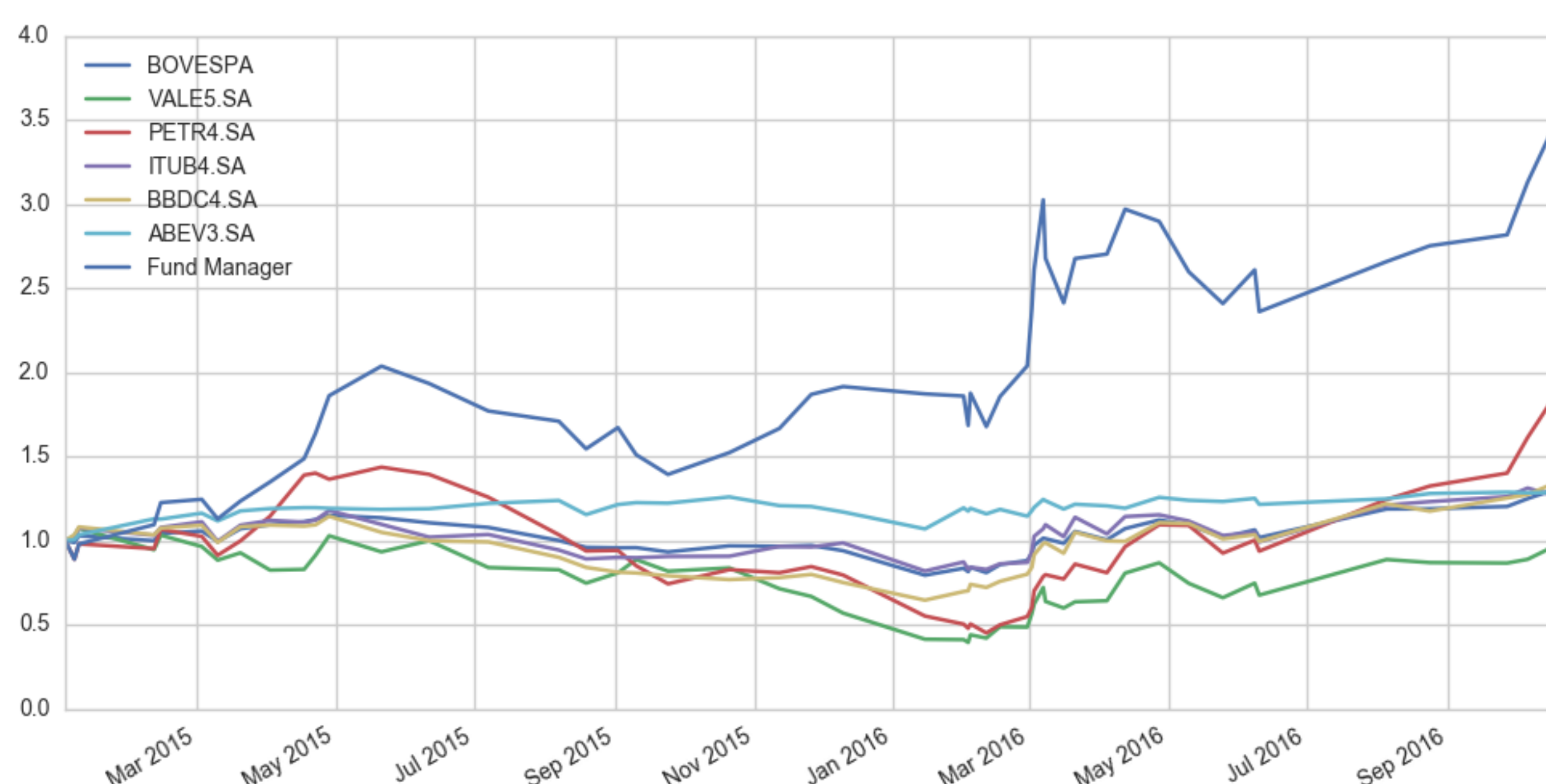
ORACLIA was implemented using several open source Python packages. Scikit-learn and tensorflow were used for machine learning algorithms and data preprocessing features. Pandas, matplotlib, seaborn and jupyter notebooks were used for data wrangling and visualization. The multiagent implementation is based on SPADE, which is a FIPA-compliant multi-agent environment based on JADE.

Conclusions and future work

The multi-agent approach of ORACLIA, where autonomous agents can operate independently in different locations, facilitates distribution and scalability. The use of the FIPA-ACL communication protocol simplifies the introduction of agents from other FIPA-compliant platforms.

Since the agents in ORACLIA's architecture reflect roles in an investment fund structure, the proposed approach could be deployed as part of the organization. Likewise, human agents could also play the role of a Strategy Agent or even the Fund Manager in the system.

The preliminary results found in the implementation of ORACLIA are promising, and can be further improved by increasing the number of assets per portfolio and the range of strategies applied for each asset. More experiments are required to ensure the model's reproducibility in different test and real-world scenarios.



Best results achieved by Oraclia, with a 252% portfolio valuation in 22 months