

Super Intelligence for The Stock Market

[Richard Craib](#)

[Follow](#)

Numerai

--

13

Listen

Share

[Numerai](#) is synthesizing machine intelligence to command the capital of an American hedge fund. Here's how.

Combining Intelligence

In the late 90s and early 2000s, new algorithms such as [AdaBoost](#) and [Random Forests](#) made breakthroughs in machine learning. The principle behind each of these algorithms was very simple: build many decision tree models that each learn something different, and then average them all together to create an ensemble model.

These algorithms worked incredibly well. They were easy to understand, and computationally efficient. Decision tree ensembles were so effective that in many cases they would outperform complicated neural networks on machine learning benchmarks. To compete, neural network researchers needed to find a way to harness their power: the power of combining many models.

In 2013, Geoffrey Hinton's team at the University of Toronto proposed a machine learning technique for neural networks called '[Dropout](#)'. Dropout was a new efficient way to ensemble the intelligence of many different neural networks, and it worked. Using Dropout to combine the intelligence of many different neural networks, Hinton's team achieved state-of-the-art performance on supervised learning tasks in vision, speech recognition, document classification and computational biology.

Invisible Hands

At [Numerai](#), we're ensembling machine intelligence from thousands of data scientists around the world to achieve breakthroughs in stock market prediction accuracy.

Every data scientist on Numerai is solving the same problem using the same underlying features. But every data scientist approaches the problem in their own unique way. With many different solutions to the same problem, Numerai is able to combine each model into a meta model just like Random Forests combines decision trees into a forest.

No data scientist on Numerai has a machine learning model that is better than all the other models combined. So Numerai is not a search for the 'best' model; it is a platform to synthesize many different models with many different characteristics. Although data scientists compete to place on [the leaderboard](#), the competition is designed to collect models. Numerai is not really a competition; it's an invisible collaboration to build the meta model.

Corollaries in Portfolio Theory

The benefits of the meta model extend beyond just increased return. Combining many different models also has remarkable consequences in portfolio theory.

Think of every model on Numerai as a biased coin (with the edge in your favor). To allocate capital in our hedge fund based on just one model would be similar to betting the entire portfolio on just one biased coin.

It is preferable from a risk standpoint to place many simultaneous bets on many independent coin flips. With enough bets, it becomes increasingly probable that the edge in your biased coin will be realized. This is illustrated by the slope of the [binomial cumulative distribution function](#).

Numerai's meta model gains simultaneous exposures to every model meaning our hedge fund holds many more independent bets than a portfolio built from just one model. The diversity of models leads to diversification in the portfolio and thereby reduces risk.

As it turns out, having lower risk is dual with higher return. In the coin flipping example, the independence of each coin flip

implies larger bet sizes from the [Kelly criterion](#). In portfolio theory, a lower risk portfolio can rationally taken on more leverage.

With many orthogonal vectors of edge, Numerai can not only earn higher returns with lower volatility but also rationally take on larger exposures.

From First Principles

Ensemble theory has a mathematical basis in both machine learning and portfolio management. Ensembling many diverse models permits lower error rates in machine learning, higher returns on individual trades, lower portfolio volatility, and higher portfolio exposure. Each of these form powerful first principles arguments for Numerai.

We are building the largest ensemble of stock market machine learning models in the world.

So far we have 191,766.