Towards Explainable Meta-Models for Ensembles of Financial Alphas

An introduction to explainable meta-learning in systematic trading

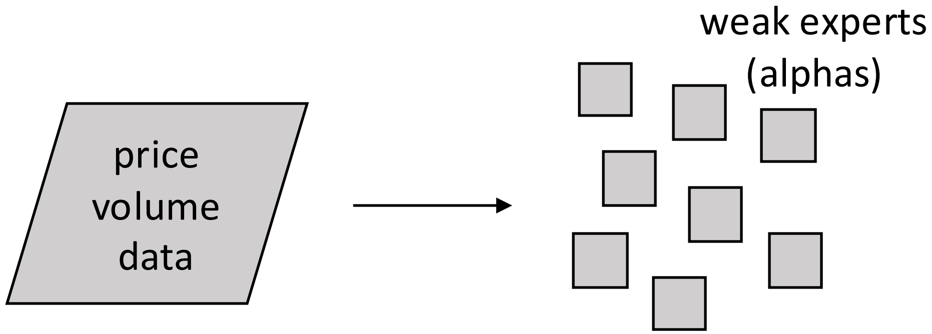
Bartłomiej Małkus

Doctoral School of Exact and Natural Sciences
Technical Computer Science

GEIST
Research Group

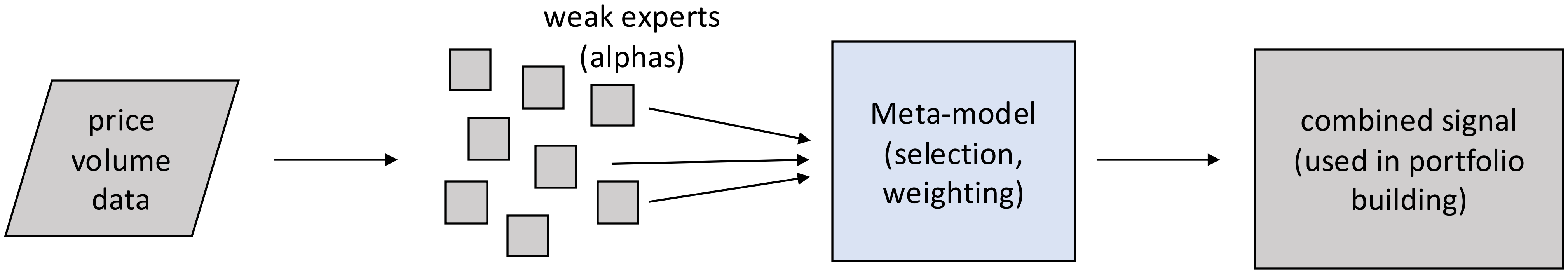
What this talk is about

- We consider systems that use many small predictive signals/models (alphas) built based on price/volume data.



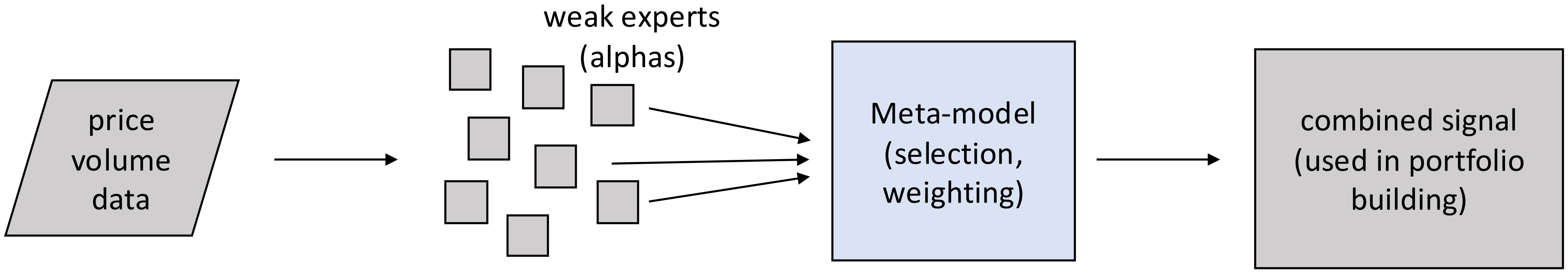
What this talk is about

- We consider systems that use many small predictive signals/models (alphas) built based on price/volume data.
- These signals are combined dynamically into an ensemble whose composition changes over time.



What this talk is about

- We consider systems that use many small predictive signals/models (alphas) built based on price/volume data.
- These signals are combined dynamically into an ensemble whose composition changes over time.
- From an ML point of view, this looks like:
  - an ensemble of weak experts,
  - a meta-model deciding which experts to trust under which conditions.
- This talk is an introduction to this setting.



Basic objects

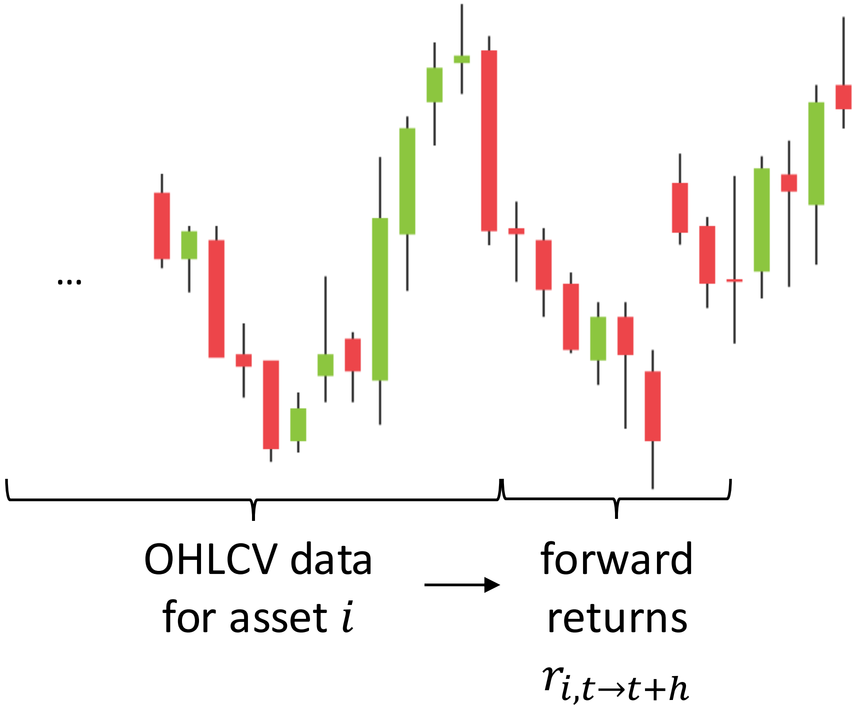
Universe of assets (e.g., stocks) $i = 1, \dots, N$

Time indexed by days t
(data can be of different granularity as well, like 1 Hour, 5 Mins, 1 Min, …)

For each (i, t):

- Daily market data
  (Open-High-Low-Close-Volume)
- Forward returns $r_{i,t \to t+h}$
  (Horizon h can be 1 day or longer)

We care only about cross-sectional ordering of returns across assets on each day.



Terminology notes

- Factor portfolio
  - long-short portfolio built from a signal to measure how “tradeable” the signal is.
- Cross-sectional vs time-series
  - Cross-sectional: compare assets to each other at a given time (ranks across stocks).
  - Time-series: follow one asset over time.
- Backtesting
  - Simulating how a signal/strategy would have performed on historical data with the usual pitfalls, like look-ahead bias, data snooping, costs, etc.
- Market regime
  - Coarse description of the current market state (e.g. {up, flat, down} x {low volatility, high volatility})

What is an alpha?

In general:

Alpha = source of return advantage (“edge”) over the market

What is an alpha?

In general:

Alpha = source of return advantage (“edge”) over the market

In our case:

Alpha = small predictive model (formula)

What is an alpha?

In general:

Alpha = source of return advantage ("edge") over the market

In our case:

Alpha = small predictive model (formula)

- it operates on OHLCV data
- it outputs a score $\alpha_{i,t}$ for each asset i at time t
- higher score ≈ higher expected future return (in cross-section)
- built from past price and volume data using time-series operators

What is an alpha?

In our case:

Alpha = small predictive model (formula)

- it operates on OHLCV data
- it outputs a score $\alpha_{i,t}$ for each asset i at time t
- higher score ≈ higher expected future return (in cross-section)
- built from past price and volume data using time-series operators

Example alpha:

$$\alpha_{i,t} = \text{ts\_mean}\left(C_{i,t}^{0.3}, 5\right) - \text{ts\_std}(V_{i,t}, 10)$$

Evaluating an alpha

Given scores $\alpha_{*,t}$ and future returns $r_{*,t\to t+h}$ compute Information Criteria

$$IC_t = corr_{spearman}(\alpha_{*,t}, r_{*,t\to t+h})$$

across assets, and built a portfolio:

long top X%, short bottom X%, equal-weight

then calculate over time:

$IC_t \to$ mean, std, IR=mean/std

portfolio → return (PnL), Sharpe, drawdowns

Challenges of applying ML in this setting

- Targets are indirect and noisy (ranks, portfolio behavior, risk).
- Very low signal-to-noise → strong risk of overfitting.
- Non-stationarity and regime changes.
- Practitioners often favor simpler, transparent models.
- Evaluation is more than test error (costs, turnover, liquidity, etc.).

Motivation

One area of quantitative finance research focuses on **large libraries of formulaic alphas**.

In practice, these alphas are **combined dynamically** into an ensemble based on their performance and sometimes market conditions.

From an ML perspective:

- we have **many weak experts** (alphas) with diverse behaviors,
- a **dynamic ensemble mechanism** deciding which ones to trust when.

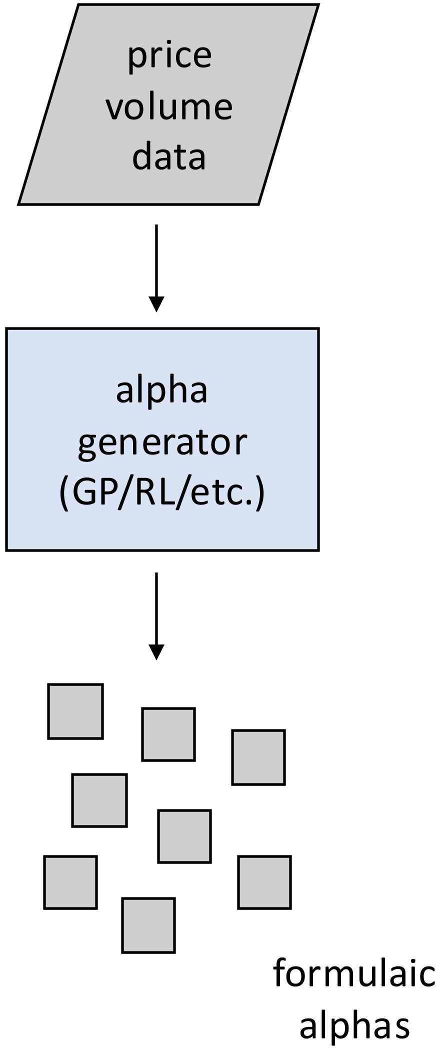
Our interest:

- characterizing and clustering the alphas,
- understanding the ensemble.

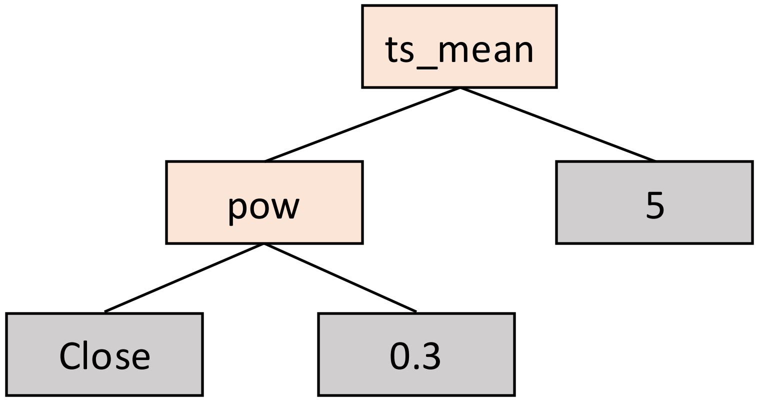
Existing Works: Formulaic Alpha Discovery

Automated alpha mining approaches:

- genetic programming/symbolic regression over expression trees,
- DL and RL over sequence/RPN representations of formulas.



expression tree
form of alpha



RPN form of alpha

Close 0.3 pow 5 ts_mean

Existing Works: Formulaic Alpha Discovery

Automated alpha mining approaches:

- genetic programming/symbolic regression over expression trees,
- DL and RL over sequence/RPN representations of formulas.

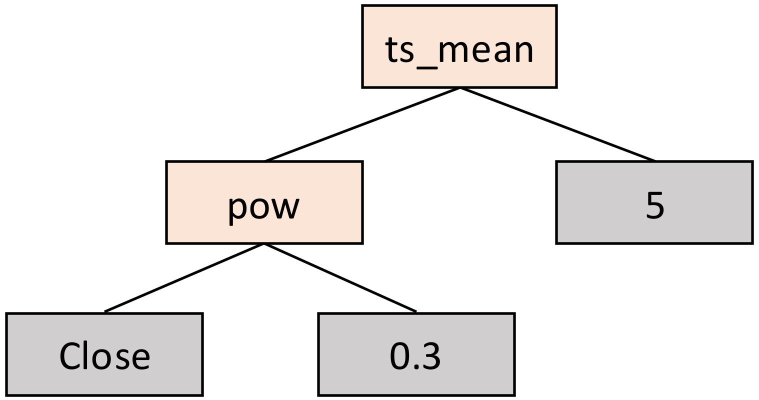
Common pattern:

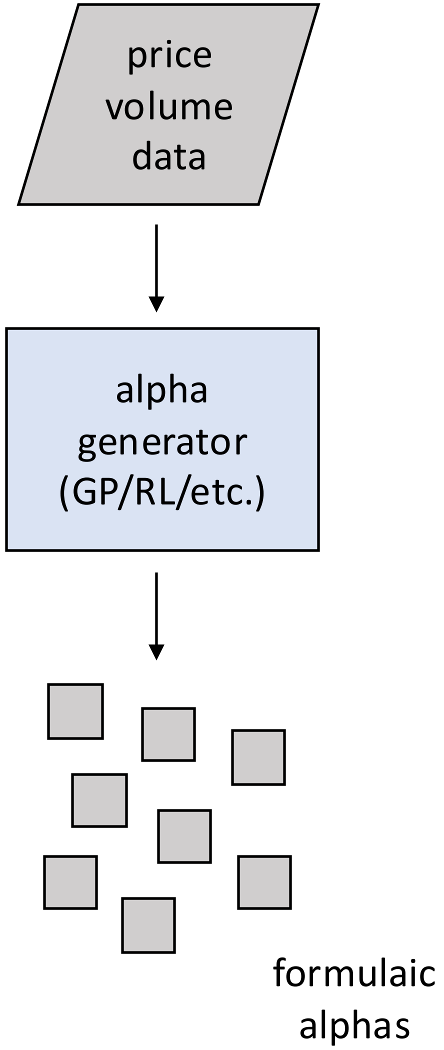
- generate many candidate alphas,
- filter by IC, ICIR, correlation, stability,
- keep a library of weak but diversified predictors for downstream combination.

RPN form of alpha

Close 0.3 pow 5 ts_mean

expression tree form of alpha



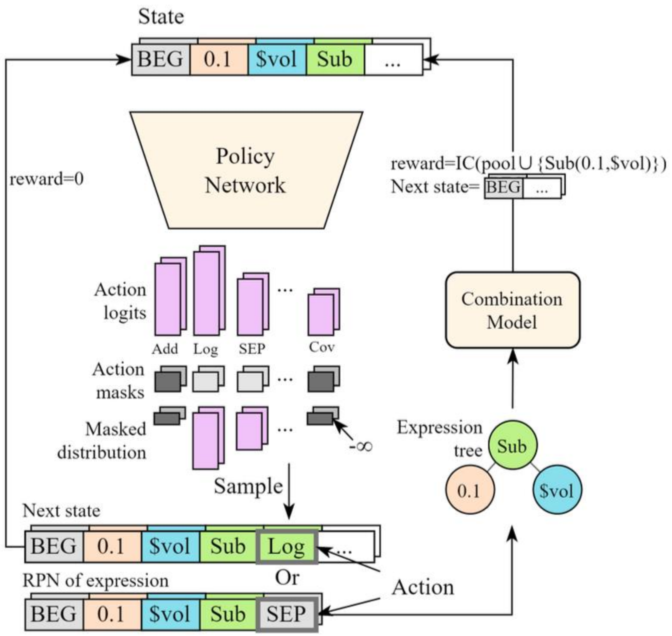


Existing Works (selected)

Idea: Treat alpha discovery as a **sequential decision problem** in formula space (RPN).

Policy builds formulas token by token.

Reward: performance of combined model that uses whole set of discovered alphas.



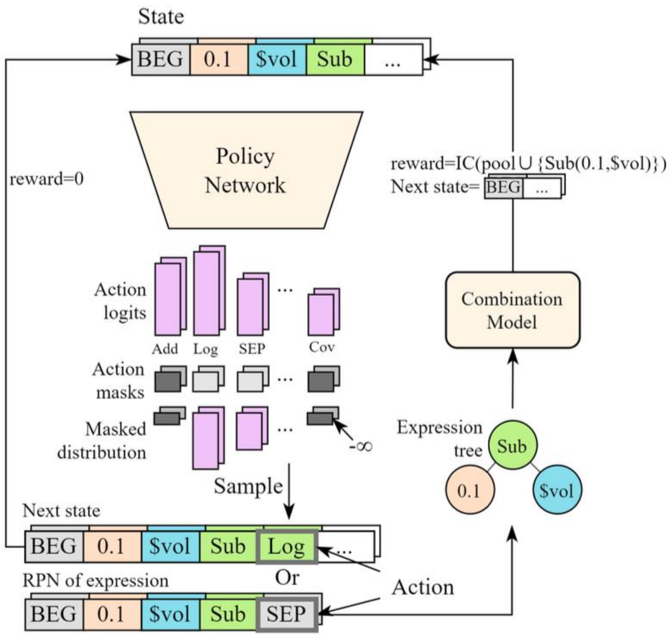
AlphaGen

Existing Works (selected)

Idea: Treat alpha discovery as a **sequential decision problem** in formula space (RPN).

Policy builds formulas token by token.

Reward: performance of combined model that uses whole set of discovered alphas.

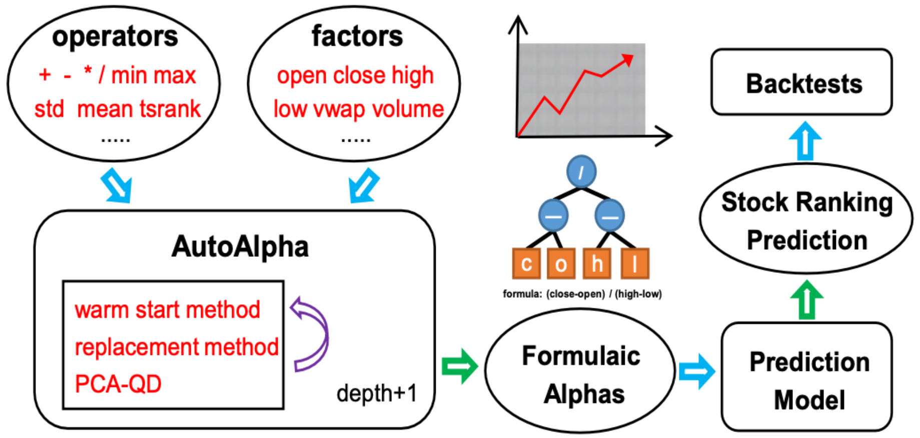


AlphaGen

Idea: Classic genetic programming, but with:

- structured search space
- explicit diversity maintenance

LightGBM/XGBoost to ensemble the results.

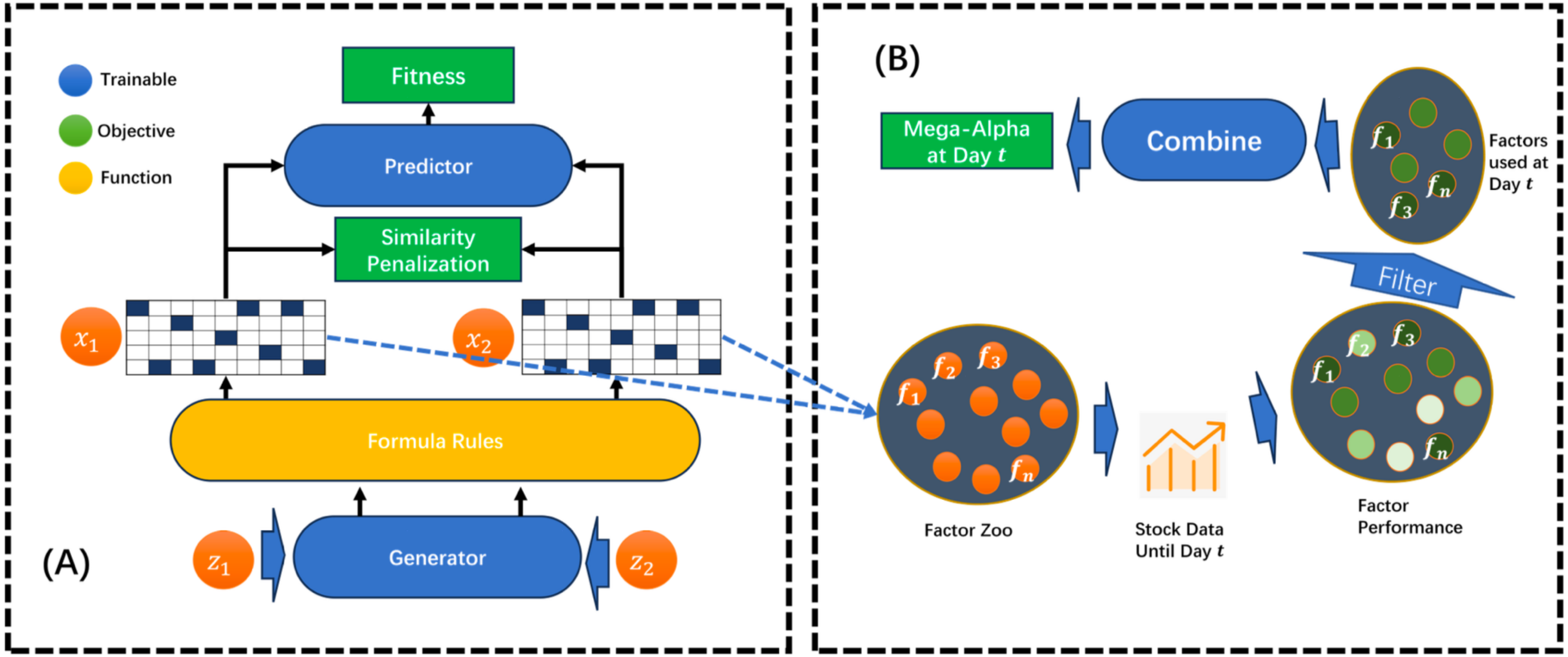


AutoAlpha

Existing Works: AlphaForge (our baseline)

Stage 1: Factor mining
- Generative network
- Fitness function network

Stage 2: Dynamic combination
- Rolling metrics for alphas
- Cross-sectional linear model

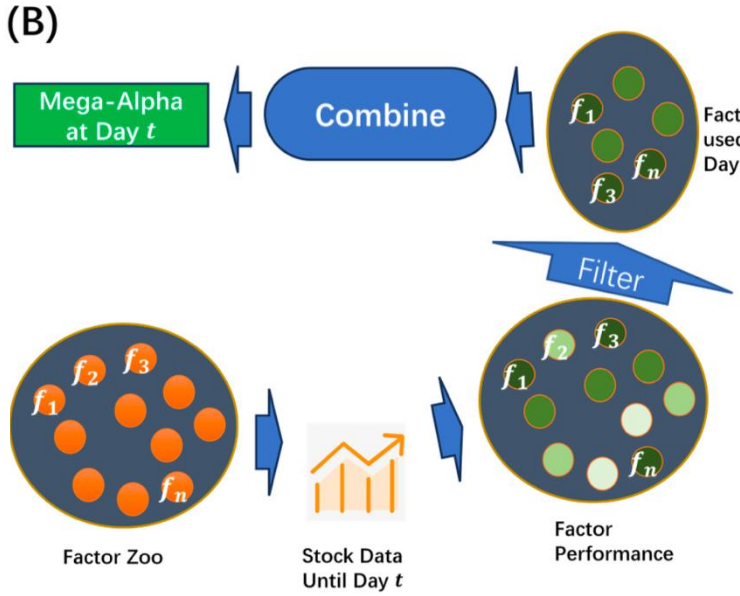


Dynamic combination baseline

- Start with a library of K alphas α_k
- For each alpha, compute rolling IC on window W
- At each day t:
  - select top-*N* alphas with IC above threshold as S_t,
  - fit a cross-sectional regression:

$$r_{i,t\to t+h} \approx \sum_{k\in S_t} \beta_{k,t}\alpha_{k,i,t}$$

(regression coefficients $\beta_{k,t}$ act as weights)



Limitations of the baseline

- Arbitrary design:
  - thresholds on IC, choice of window W, top-*N* are arbitrary and static.
- Instability:
  - selected set S_t and weights $\beta_{k,t}$ can vary strongly over time,
  - it's hard to distinguish meaningful shifts from noise.
- No modeling of structure or context:
  - each alpha is treated as an index, not as a model with features,
  - decisions do not account for market conditions.
- Explainability is superficial (if any):
  - answers "*which alphas are important today?*", but does not answer "*why they are important?*"

Alpha features

- We associate each alpha with:
  - structural features (what the formula looks like),
  - behavioral features (how it tends to behave).

max window
(short/mid/long term)

$$\mathrm{IC} = \cdots$$
$$\mu_W(\mathrm{IC}) = \cdots$$
$$\sigma_W(\mathrm{IC}) = \cdots$$
$$\mathrm{PnL}_{portfolio} = \cdots$$
...

$$\alpha_{k,i,t} = \mathrm{ts\_mean}\left(C_{i,t}^{0.3}, 5\right) - \mathrm{ts\_std}(V_{i,t}, 10)$$

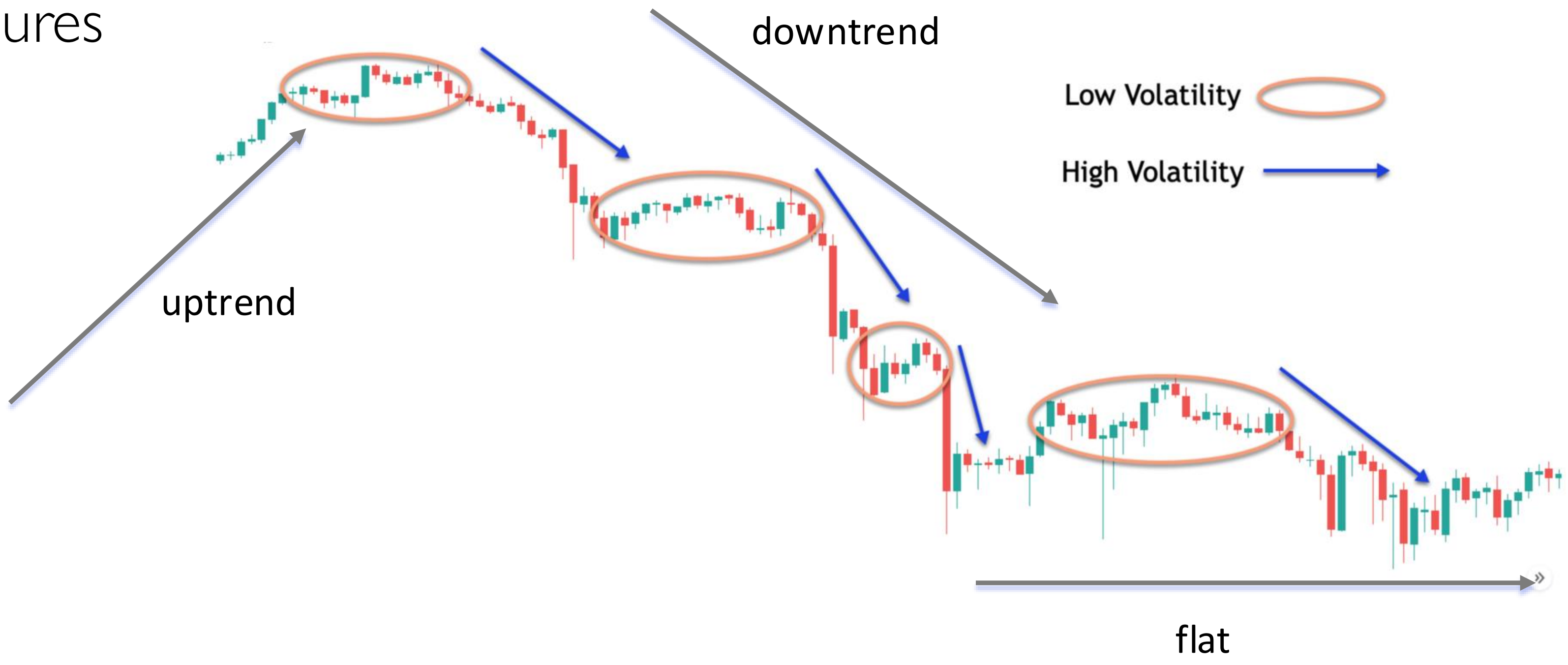
tree depth
complexity
(based on AST)
...

uses smoothing

uses volatility

Example market features / regimes

- We derive simple market features, for example:
  - trend x volatility, 6 categories (regimes):
    {up, flat, down} x {low volatility, high volatility}
  - cross-sectional dispersion of returns
  - liquidity-related features

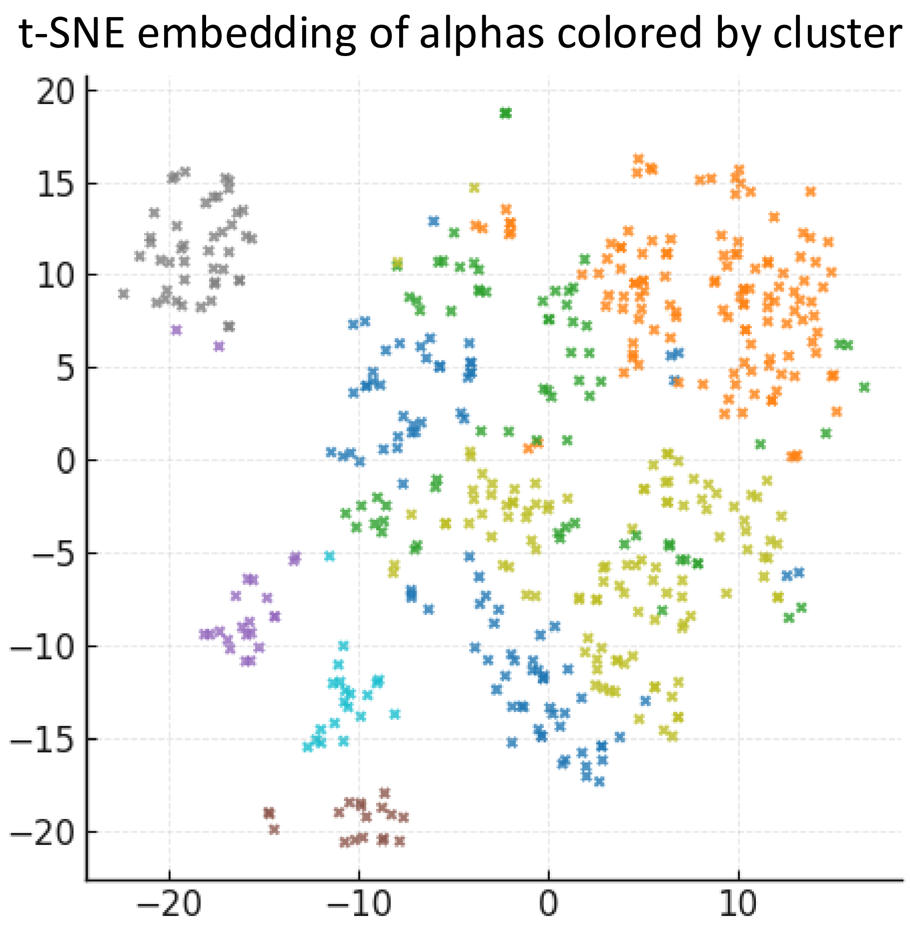


Combining problem formulation (informal)

- Each alpha is a small model $f_k(\cdot)$ producing scores $\alpha_{k,i,t}$.
- We associate each alpha with:
  - structural features (what the formula looks like),
  - behavioral features (how it tends to behave).
- We derive simple market features.
- Our meta-model (that we propose to build):
  - operates on alpha and market features,
  - aims to characterize combiner decision patterns,
  - provides explanations which kind of alphas are used under which market conditions.

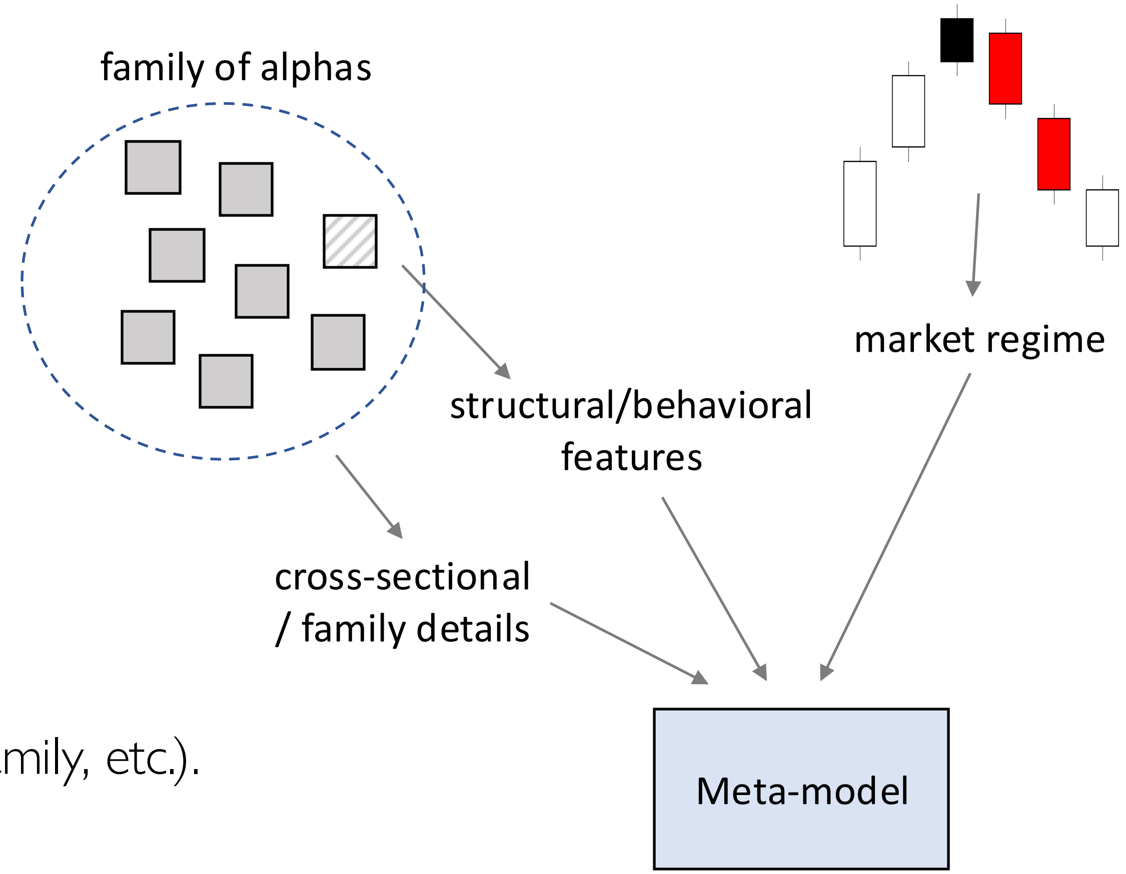
Clustering alphas into families

- We build a feature vector for each alpha from structural and behavioral features.
- We apply clustering to obtain families of alphas:
    - global families of alphas (often matching intuitive types, like short-term mean-reversion, momentum, volume-based, etc.),
    - optionally, regime-specific families.
- Families can be used:
    - as higher-level units of analysis,
    - for aggregating meta-model decisions,
    - as a guide for future alpha generation / search.



Meta-learning setup

- Training examples are pairs (α_k, t).
- Feature vector $x_{k,t}$ include s:
  - structural features of α_k,
  - behavioral features of α_k for time window W,
  - family label,
  - market regime at t
  - details on cross-section of alphas (like rank in family, etc.).
- Target $y_{k,t}$ can be chosen as:
  - selection indicator (whether α_k is used at t),
  - weight $\beta_{k,t}$,
  - future usefulness (IC/portfolio Sharpe ratio over a forward window).
- We then learn a meta-model $g_\theta : x_{k,t} \mapsto y_{k,t}$ as an interpretable model and analyze it using XAI tools.



Explaining the meta-model

- Once the meta-model is fitted, we can analyse it using XAI tools:
  - global feature importance,
  - dependence analysis for selected features,
  - stability of explanations across time and regimes.
- Example questions:
  - *Do short-horizon alphas become more important in certain volatility regimes?*
  - *Are volume-based signals preferred in specific conditions?*
  - *Are meta-model explanations robust, or do they follow the noise?*

Summary

- Formulaic alphas can be viewed as small predictive models / weak experts.
- Dynamic combiners from the literature produce time-varying selection and weighting of these experts, but they lack true interpretability.
- We can construct a meta-dataset by deriving features from alphas and the market, and clustering alphas into families.
- An explainable meta-model over this dataset allows us to:
  - identify ensemble decision patterns,
  - study how different alpha families are used across regimes,
  - investigate stability vs noise in time.

Open questions

- How rich should structural and behavioral feature sets be?
- How to handle non-stationarity and model regime shifts systematically?
- How to distinguish true factor timing from overfitting to noise?
- How stable should explanations be to be actionable?
- Are there analogous settings beyond finance we could test?