

# ML in Finances: Homework 2

## Пункт 1

Задача сводится к бинарной классификации направления движения цены на следующем шаге. Пусть  $r_{t,t+1}$  — доходность актива. Определим целевую переменную:

$$Y_{t+1} = \mathbb{I}[r_{t,t+1} > 0]$$

В качестве бенчмарка используется:

$$w_t^{Bench} = \text{sign}(r_{t-12,t})$$

где  $r_{t-12,t}$  — доходность за прошлые 12 месяцев.

В предлагаемом методе позиция определяется вероятностью, полученной из CART:

$$w_t^{CART} = \begin{cases} 1, & \hat{P}(Y_{t+1} = 1|X_t) > 0.5 \\ -1, & \hat{P}(Y_{t+1} = 1|X_t) \leq 0.5 \end{cases}$$

## Признаки (Features)

Вектор признаков  $X_t \in \mathbb{R}^3$  включает:

1.  $MOM_{12}$ : Доходность за 12 месяцев (с лагом 1 месяц)  $\frac{P_{t-1}}{P_{t-12}} - 1$ .
2.  $MOM_1$ : Доходность за последний месяц  $\frac{P_t}{P_{t-1}} - 1$ .
3.  $VOL$ : Годовая волатильность  $\sigma_{252} \cdot \sqrt{252}$ .

## Алгоритм тестирования (Expanding Window)

Используем схему расширяющегося окна, чтобы избежать подглядываний и, соответственно, утечек данных.

### Псевдокод алгоритма

*Инициализация:*

- начальная дата старта  $T_{start}$  (2015-01-01);
- начальный набор данных  $D_0$ .

$t$ -й шаг (ребалансировка каждый месяц):

1. сформировать  $D_{train} = \{(X_i, Y_i)\}_{i=1}^t$ ;
2. обучить DecisionTreeClassifier на  $D_{train}$ ;

3. получить прогноз  $\hat{p}_t$  для тестовой точки  $X_{t+1}$ ;
4. сформировать позицию  $w_t$  и рассчитать  $R_{strat} = w_t \cdot r_{real}$ ;
5. расширить выборку:  $D_{train} \leftarrow D_{train} \cup (X_{t+1}, Y_{t+1})$ .

*Гиперпараметры:* В модели использую:  $\max_{depth} = 4, \min_{samples\ leaf} = 50$ .

Глубину ограничил 4, так как финансовые временные ряды зашумлены, и глубокое дерево приводит к переобучению.

## Пункт 2

Результаты тестирования Out-of-Sample (2015-2025).

Таблица 1: Summary Statistics: CART vs Benchmark

Metric	Benchmark (TSMOM)	CART Strategy
Avg. Return	0.07%	0.50%
Return Max	4.20%	8.20%
Return Min	-13.31%	-18.72%
Ann. Return	0.54%	5.41%
Ann. Vol.	7.61%	11.90%
Sharpe Ratio	0.07	0.46
Info. Ratio	-	0.38
Kurtosis	10.02	6.55
Skewness	-2.12	-1.32
MDD	-28.48%	-26.98%

### Интерпретация результатов:

- **Доходность:** CART существенно превосходит бенчмарк по Ann. Return (5.41% против 0.54%). Средняя доходность выросла за счет фильтрации ложных сигналов в периоды высокой волатильности.
- **Риск:** Sharpe Ratio увеличился с 0.07 до 0.46. Несмотря на рост Volatility (с 7.61% до 11.90%), качество доходности на единицу риска улучшилось. Information Ratio = 0.38 указывает на наличие альфы.
- **Хвосты распределения:** Kurtosis у CART (6.55) ниже, чем у Benchmark (10.02). Это важный результат: стратегия на базе ML реже ловит "черных лебедей" (momentum crashes), чем простой TSMOM. Skewness также улучшился (стал менее отрицательным).

## Пункт 3

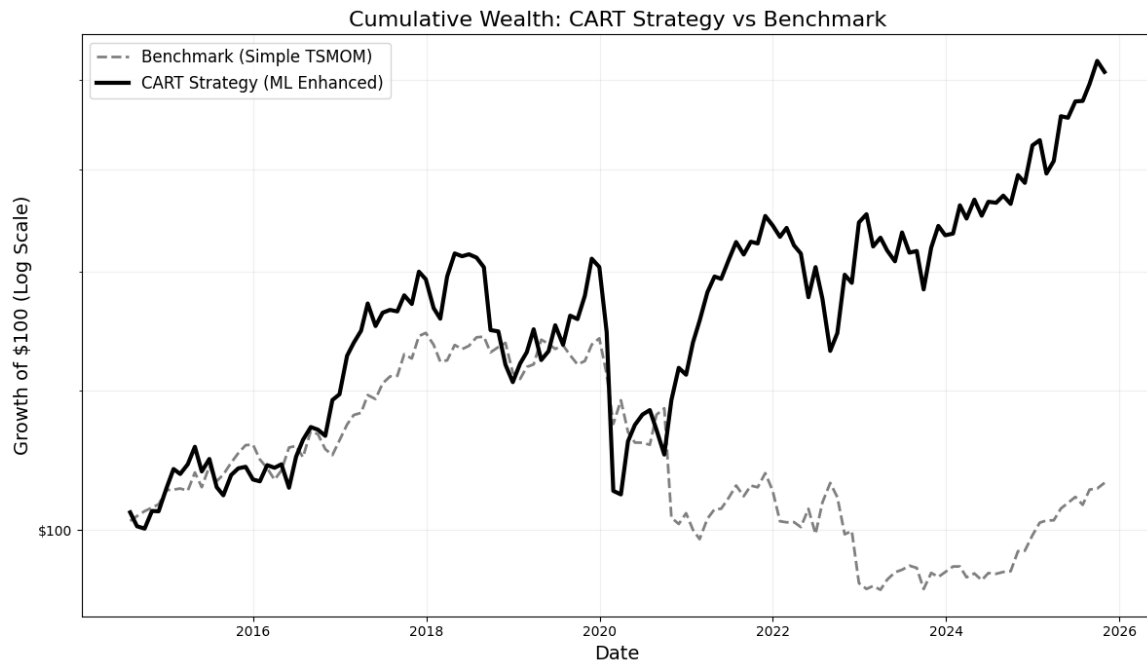


Рис. 1: Cumulative Wealth (Log Scale). Comparison of growth of \$100 invested in Benchmark vs CART Strategy.

График в логарифмической шкале показывает, что CART обеспечивает более стабильный наклон кривой капитала на периоде 2015-2025, в то время как бенчмарк показывает стагнацию.

## Пункт 4

Дополнительно можно отметить, что MDD (максимальная просадка) у CART ниже (-26.98% против -28.48%), что согласуется с теоретическим ожиданием: дерево перестает торговать (или меняет позицию), когда паттерны волатильности становятся аномальными, тогда как линейный TSMOM продолжает удерживать позицию.