Chapter5: Financial Labels

Ali Abedi and Ehsan Tabatabaei

July 2020

Motivation

- Classification vs regression problems,
- Why are financial labels important?

Fixed-horizon method

- $r_{t_{i,0},t_{i,1}} = \frac{p_{t_{i,1}}}{p_{t_{i,0}}} - 1,$
- $y_i = \begin{cases} -1 \text{ if } r_{t_{i,0},t_{i,1}} < -\tau, \\ 0 \text{ if } \left| r_{i_{i,0},t_{i,1}} \right| \leq \tau, \\ 1 \text{ if } r_{t_{i,0},t_{i,1}} > \tau, \end{cases}$

  where $r_{t_{i,0},t_{i,1}}$ is the change percentage of i^{th} feature from time t_0 to t_1 and τ is a fixed threshold.

Fixed-horizon method problems

- Financial data has heteroscedasticity,
- Dismisses intermediate returns,
- Investors do not forecast returns for an exact period.

Solutions for the heteroscedasticity problem:

- Applying fixed-horizon on tick bars,
- Using Standardised returns:
- $z_{t_{i,0},t_{i,1}} = \frac{r_{t_{i,0},t_{i,1}} - \mu}{\sigma}$
- $y_i = \begin{cases} -1 \text{ if } z_{t_{i,0},t_{i,1}} < -\tau, \\ 0 \text{ if } \left| z_{i_{i,0},t_{i,1}} \right| \leq \tau, \\ 1 \text{ if } z_{t_{i,0},t_{i,1}} > \tau, \end{cases}$

For the other two, de Prado suggests the next methods...

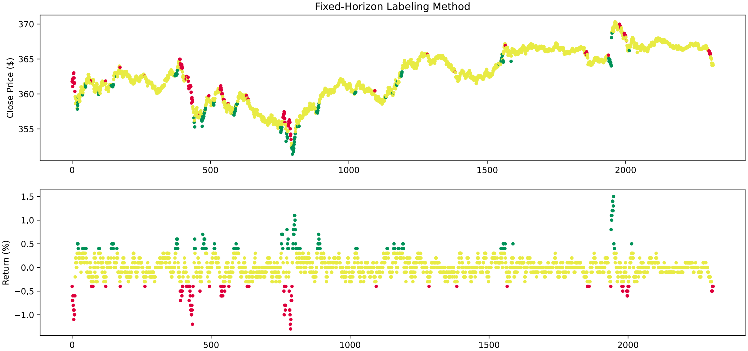

Figure 1: Fixed-horizon labeling raw returns

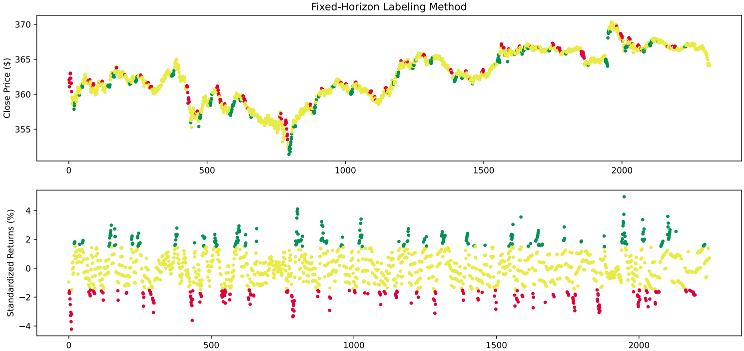

Figure 2: Fixed-horizon labeling standardized returns

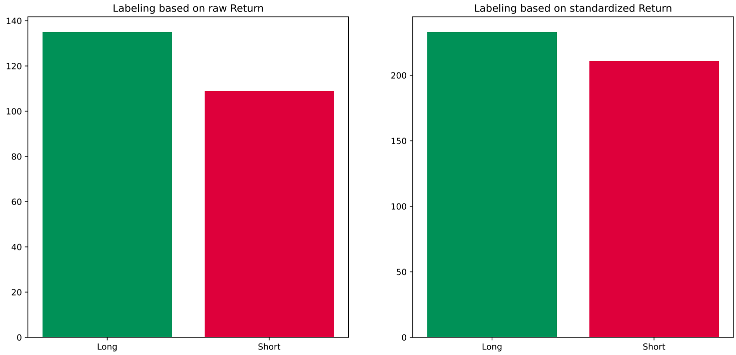

Figure 3: Labels comparison between raw and standardized returns

Triple-Barrier method

A realistic method of how do asset managers really act. Holding a position can end to one of the below:

1. profit target is achieved,
2. stop loss limit is reached,
3. the position is closed after certain bars.

Thus, if we set a profit target, a stop loss limit, and a maximum holding period, we can label the data as +1, -1, 0 respectively. (or $sign(r_{t_{i,0},t_{i,1}})$).

Triple-Barrier method

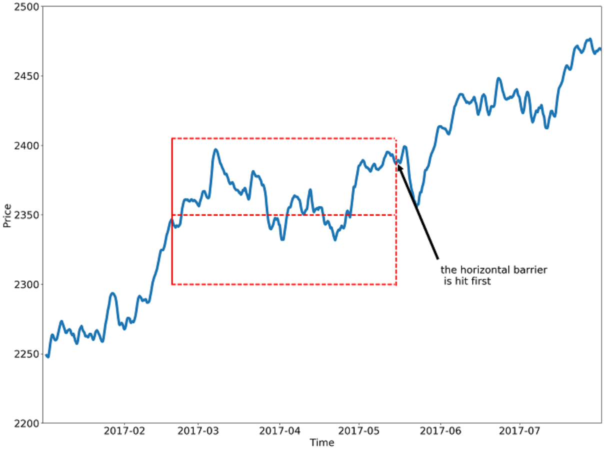

Figure 4: Triple-barrier horizon

Triple-Barrier method problems

- Maybe position side is unknown,
- Setting 3 parameters changes the results a lot,
- Touching a barrier is a discrete event.

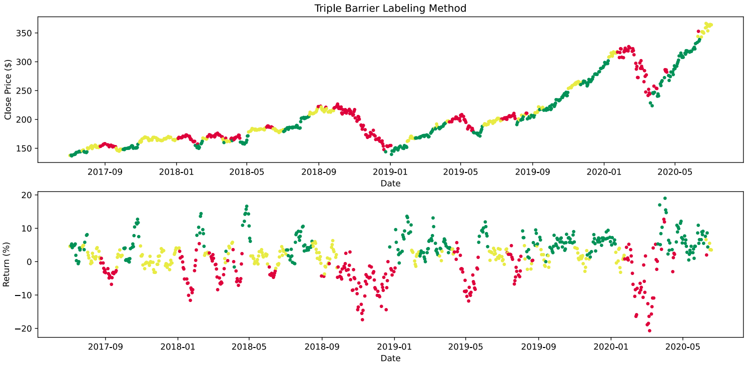

Figure 5: Triple-Barrier method

Trend-scanning method

What constitutes a trend?

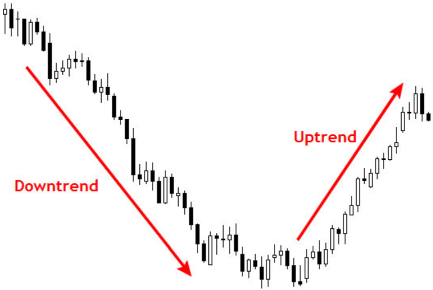

Figure 6: Upward vs Downward Trend

How to implement?

$$x_{t+l} = \beta_0 + \beta_1 l + \varepsilon_{t+l},$$

$$\hat{t}_{\hat{\beta}_1} = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}1}},$$

- L is chosen in a way that, $\hat{t}_{\hat{\beta}_1}$ is maximized
- The sign of β coefficient used as the label.

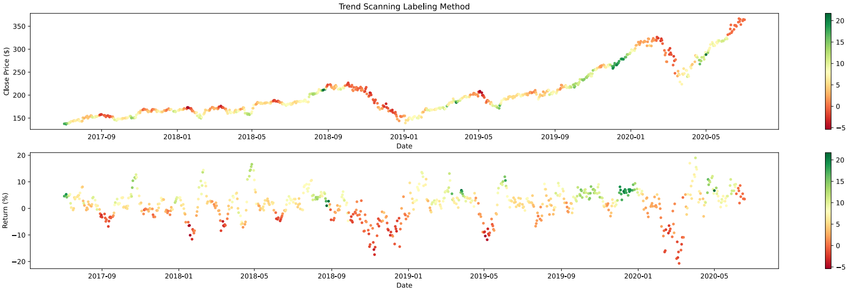

Figure 7: Trend-Scanning Method

The end of the first section...

Meta-labeling

Using meta-labeling has many points including:

- Turning a weak predictor into a strong predictor,
- Enables building ML models on white-boxes,
- Most importantly, can be used for bet-sizing calculation.

Model performance

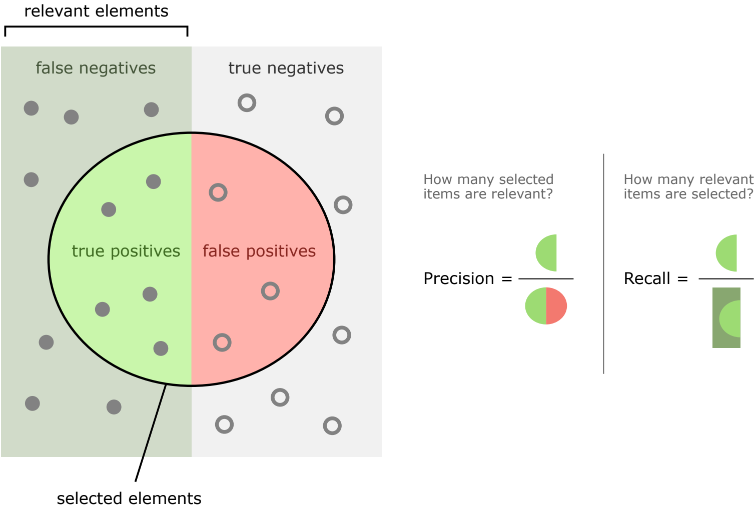

Figure 8: Confusion matrix

Secondary Model

- the primary model predicts the position side,
- the secondary model predicts the profitability of the primary model,
- The final label is the multiplication of the two,
- This reduces recall and increases precision,
- Overall, it results in a better F1.

Bet Sizing

After deciding on the labeling, next we want to discuss about size of positions:

- Position side vs position size,
- The effect of bet sizing on our return,
- Decision on the size based on the model performance.

Bet Sizing by Expected Sharpe Ratio

Let p be the expected probability that the opportunity yields a profit π,and $1-p$ the expected probability that the opportunity yields a profit $-\pi$. The expected return, under Bernoulli distribution assumption, is:

$$\mu = p\pi + (1-p)(-\pi) = \pi(2p-1),$$

and the Sharpe ratio defined as

$$z = \frac{\mu}{\sigma} = \frac{p - \frac{1}{2}}{\sqrt{p(1-p)}},$$

Assuming that the Sharpe ratio follows a normal standard distribution:

$$m = 2Z[z] - 1.$$

Ensemble Bet Sizing

Consider having multiple (n) meta-labeling classifiers (each of them is from a Bernoulli) then the probability of having profitable position is drawn from a binomial distribution. Therefore:

$$y_i = \{0, 1\}, i = 1, \ldots, n.$$

$$\sum_{i=1}^{n} y_i \sim B[n, p].$$

Moivre–Laplace theorem

This theorem states that as $n \to \infty$ Bernoulli distribution converges to a normal distribution with mean np and variance $np(1-p)$, assuming that the predictions are i.i.d. Accordingly

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n},$$

$$y_i \sim N[p, p(1-p)/n].$$

Based upon Moivre-Laplace theorem, the average and standard deviation of n meta-labeling classifiers are

$$\bar{p} = 1/n \sum_{i=1}^{n} y_i,$$

$$\sigma(\hat{p}) = \sqrt{\hat{p}(1-\hat{p})/n},$$

and subject to the null hypothesis $H_0 : p = 1/2$ we can have

$$t = (\hat{p} - 1/2)/\sqrt{\hat{p}(1-\hat{p})}\sqrt{n},$$

Accordingly we have bet size as

$$m = 2t_{n-1}[t] - 1.$$

Conclusion

- We now know that labeling our data effects the outcome substantially,
- Understood Fixed-horizon labeling and its shortcomings,
- Found out that standard returns fixes some of them,
- Figured out how to implement triple-barrier method,
- discussed about trend-scanning method and its implementation,
- Learned about meta-labeling approach to reduce false positives,
- Got familiar with bet-sizing methods based on the model performance.

Thank you!