Time Series Prediction

Market Regime Classification

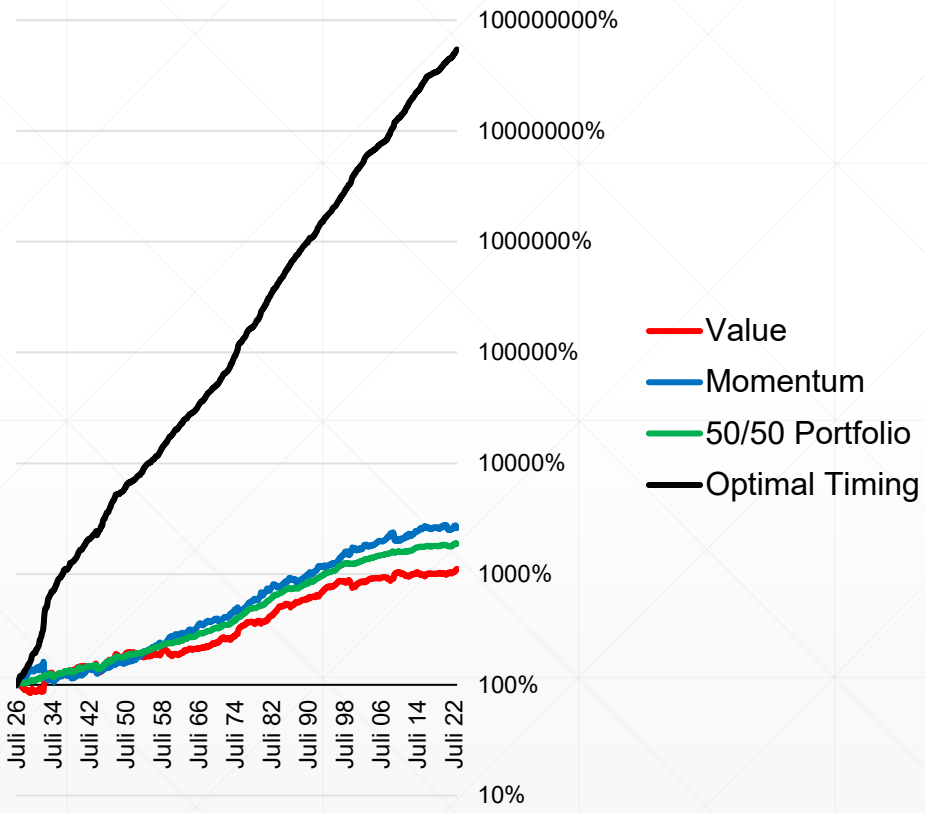
Project Milestone

Agenda

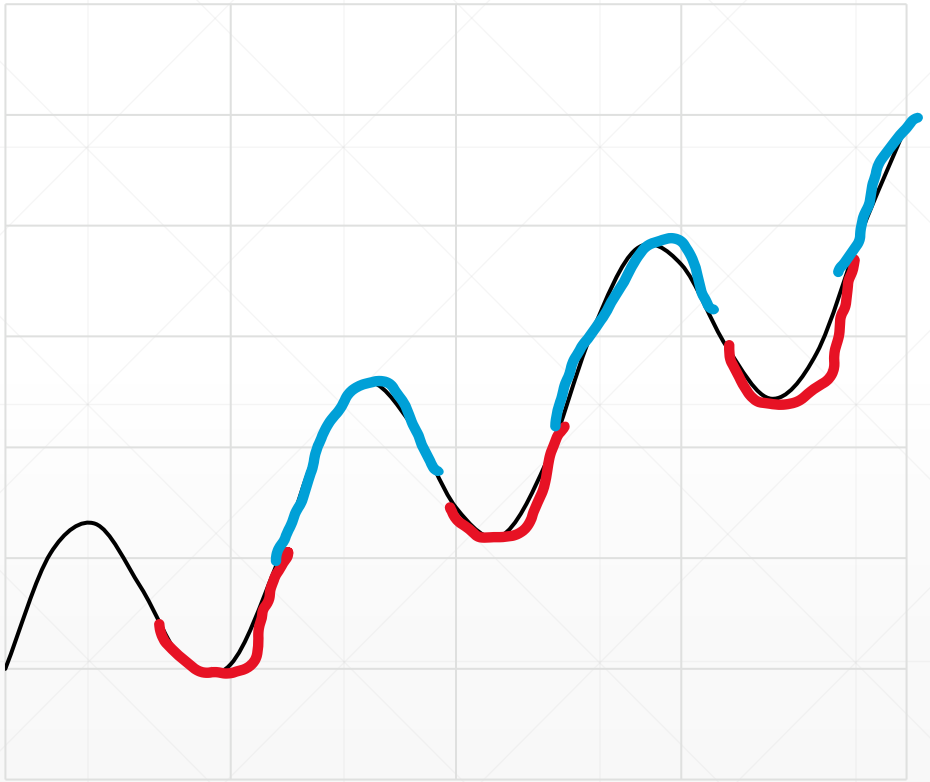
- Problem
- Value vs. Momentum
- Binary Encoding of Target Variable
- Features
- Problems
- Models

Problem

- Investors employ diverse strategies to capitalize on various opportunities and anomalies in the financial markets.
- The two predominant ones:
  - Value Investing
  - Momentum Investing
- However, both experience extended periods of underperformance, which challenge investors' confidence.
- Can we time those periods?

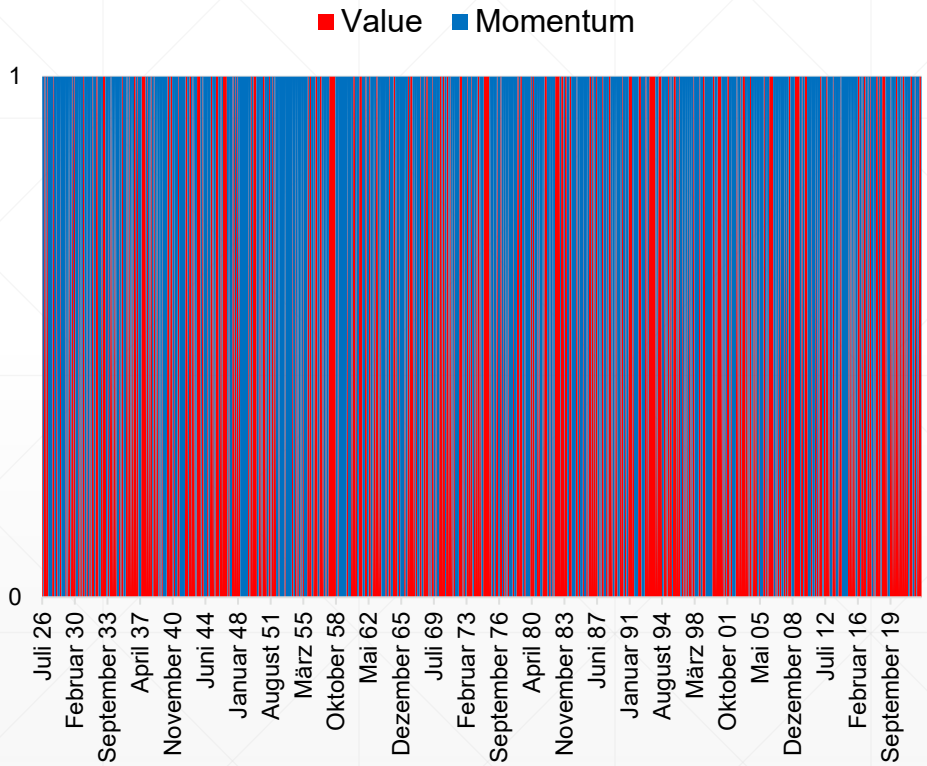


Value vs. Momentum



- Value Investing:
  - Buying undervalued securities
  - Selling overvalued securities
  - Counter-cyclical, going against prevailing market trends
- Momentum Investing:
  - Buying securities with strong performance
  - Selling securities with weak performance
  - pro-cyclical, aligning with current market trends
- Result: pronounced negative correlation (-0.5) & consequently a lot of added value through timing

Binary Enconding of the Target variable



- We want to predict whether Value or Momentum will be more profitable in the coming month $t + 1$.
- Monthly return data from Ilmanen et al. (2019); ranging from 1926 to 2023
- $Y_{t+1} = \begin{cases} 0, \text{if Value} > \text{Momentum} \\ 1, \text{if Momentum} > \text{Value} \end{cases}$
- $X_t = \{\text{some features}\}$

Features

Own Features

- Factors themselves show momentum (Gupta/Kelly 2019).
  - $v\_mom_t = \prod_{t-d}^{t}(1 + value_t) - 1$
  - $m\_mom_t = \prod_{t-d}^{t}(1 + momentum_t) - 1$
  - $d = \text{months}$
- Other price-based indicators:
  - Rolling volatility
  - Etc.

Economic Features

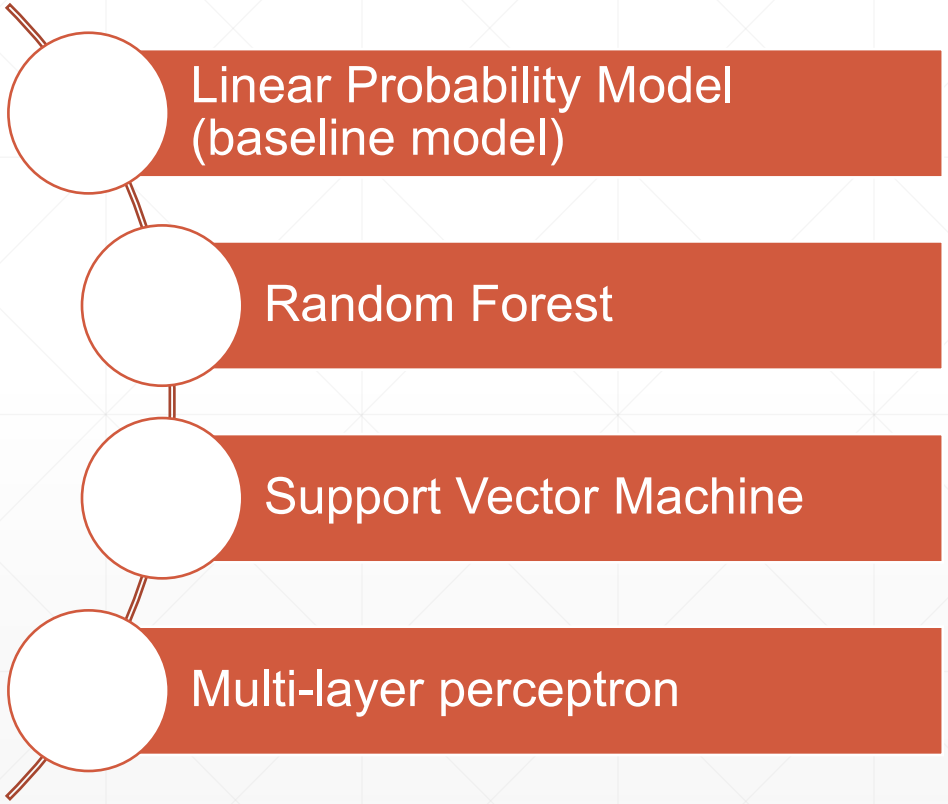
- Real Gross Domestic Product
- Unemployment Rate
- Personal Consumption Expenditures
- Effective Federal Funds Rate
- Sentiment
- Etc.

Problems

- Not all features have such a long history. There are long non-random blocks, where some features are not available.
- Some models can handle missing data automatically, like trees.
- Others not.
- Possible solution: Missingness variables $z = \begin{cases} 1, & \text{if data is missing} \\ 0, & \text{otherwise} \end{cases}$

- A lot of features will be highly correlated, especially the economic ones.
- First step is to do dimensionality reduction, like PCA

Models



- Fernández-Delgado et al. (2014) conclude, that from hundreds of classifiers, those are the best ones, especially for binary classification
- Linear probability model is a simple linear regression $Y = X\beta + U$, where $Y = \begin{cases} 1 \\ 0 \end{cases}$. The simplest binary classification model.