

Unlike the main competition, the data in the historical file isn't essential to use. The target you submit (a continuous value centered around 0.5) is based on the expected return between day 2 and day 6 after the Friday date, and you can use any data you like to come up with a signal. If you plan to derive from traditional OHLC/OHLCV market data, you might use a data source such as y-finance, and train/validate over any segments of the data you wish. Submitting values derived directly from some data without involving a model could also be effective, e.g. if you had a source of short term sentiment data, simply centering and scaling that might suffice.