**Purpose / Topic:**

**We want to exam if machine learning (ML) models could improve the returns of using technical indicators (technical trading strategies) to trade stocks.**

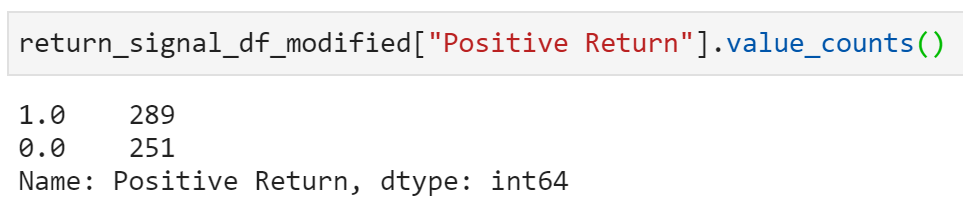
**Data:**

**In order to include the sentiment indicator, multiple sources of news have been examined, such as NEWS API, Returns. And we have made the decision to choose 2019-1-1 to current as the time frame for downloading stock data from the YAHOO FINANCE, and news data from Financialmodelingprep for this project. As the maximum free news are only available for this period.**

**Data Cleaning:**

1. **Stock data from YAHOO FINANCE is clean.**
2. **For the News data, we have removed characters that are not letters, punctuations, and stop words. The cleaned title and text have been consolidated into “content” columns to be further analyzed via Vader.**

**Check Imbalanced Data for Stock Returns：**



**Thanks. XXX for the great introduction of these ML models.**

**Firstly, let’s look at the accumulated returns for each individual technical indicator.**

**We have tested them on three stocks. Activision Blizzard Entertainment, ATVI, EBAY and TSLA.**

**The blue line with little dots is the accumulated returns for the stock over the testing period.**

**And the others lines show the accumulated returns for each indicator.**

**So, for example, EBAY, it is clear that by using one indicator alone, it is highly unlikely, able to generate returns that is at least good as simply buy and hold strategy.**

**AND for TSLA, if you could not see the blue line is because it’s overlapped with other some indicators. Because TESLA has been in a very strong upward momentum, these indicators just reflected on such strong trend.**

**Whereas Bollinger band 费奥德failed, as strategy is about when the prices cut its upper band, suggesting a SELL Signal, but the fact is the TESLA stock price was hardly able to go down, because of the strong buying momentum.**

**AND another interesting find here, is the sentiment trading strategy works so well for Blizzard, I am guessing blizzard entertainment’s share price is a little bit sensitive to news announcements.**

**SO, instead of just relying on one particular indicator, we have combined five technical indicators.**

**Because these indicators have low correlations as described early, combing them might provide a more solid signal. Just like the ML weak learners, combing them could be a strong leaner.**

**However, they might not able to give consistent signals at single point in time, so we will use ML models to help us with the trading decision.**

**By combining five indicators and with the assistance of ML models, the trading decision could be improved. Which can be seen from the graphs below.**

**For example, EBAY,**

**The blue line with little dots is the accumulated returns for the share over the testing period.**

**it is clear, using ML models could outperform the buy-hold stock strategy.**

**Especially, the Gradient Boost model, outperforms other three ML models.**

**With an accuracy about 60%.**

**AND ALSO, by comparing `F1-Scores`, `AUPRC`, and `Accuracy Scores`; `Gradient Boost Model` is recommended.**

**As it has the highest F1 scores in predicting rising / decreasing share prices, and also largest area under precision and recall curve.**

**However, it is worth noting that, for different ticker (stock), the result could change.**

**For example, TESLA (TSLA) and Blizzard (ATVI) show a different result.**

**On the left, for blizzard, only the Gradient Boost model, outperforms buy-hold stock strategy.**

**For TESLA, the overlapped curve might be caused by the inefficiency OR failure of the indicators.**

**This means that if we have more time, we will need to try different combination of indicators. And also test a variety of hyperparameters.**

**So, to conclude,**

**Technical indicators by themselves usually underperforms the buy-hold stock trading strategy. As we can see from the graphs.**

**To make decision on the combined signals, ML models could be used to significantly improve the accumulated returns. (For example, EBAY.)**

**Especially, the gradient boost model would be recommended among other models.**