Marcos Lopez De Prado firstly mentioned the idea of meta-labeling in the book Advances in Financial Machine Learning, which promises to improve strategy performance metrics, particularly helpful in enhancing F1-score by helping to filter out false positives.

The following section was directly from the textbook to ensure that interpretations were introduced without mistakes.

In statistics, binary classification problems represent a trade-off between false positives (type I. errors) and false negatives (type II errors). Generally, increasing the false-positive rate of a binary classifier tends to increase its true-positive rate. The receiver operating characteristic (ROC) curve of a binary classification problem measures the cost of increasing the true positive rate in terms of accepting higher false-positive rates.

Figure 4 illustrates the so-called “confusion matrix,” where the TP area contains the true positives, and the TN area has the true negatives. Two types of errors also included: false positives (FP) and false negatives (FN). More definitions of the confusion matrix were discussed in section A.

In general, decreasing the FP area comes at the cost of increasing the FN area because higher precision typically means fewer calls, hence the lower recall. F1-score is a harmonic combination of precision and recall, which maximizes the overall efficiency of the classifier.

Meta labeling would be extremely helpful if we want to achieve higher F1 scores. First, building a model which could achieve high recall even if the precision is not exceptionally high. Then, correcting the low precision by applying meta-labeling to the positives predicted by the primary model.

As a result, the F1 score would be increased by filtering out the false positives, where most of the majority of positives have already been identified by the primary model.

There are many reasons to illustrate that meta-labeling is a very powerful tool in machine learning areas. First, meta labeling allows you to build an ML system on top of a white box (like a fundamental model founded on economic theory). This ability to transform a fundamental model into an ML model should make meta labeling particularly useful to “quantamental” firms. Second, the effects of overfitting are limited when you apply meta labeling because ML will not decide the side of your bet, only the size. Third, by decoupling the side prediction from the size prediction, meta labeling enables sophisticated strategy structures. For instance, consider that the features driving a rally may differ from the features driving a sell-off. In that case, you may want to develop an ML strategy exclusively for long positions, based on the buy recommendations of a primary model, and an ML strategy exclusively for short positions, based on the sell recommendations of an entirely different primary model. Finally, achieving high accuracy on small bets and low accuracy on large bets will ruin you. As important as identifying good opportunities is to size them properly, it makes sense to develop an ML algorithm solely focused on getting that critical decision (sizing) right. meta-labeling ML models can deliver more robust and reliable outcomes than standard labeling models.