The Machine Learning Module

Introduction

The machine learning module within Walk Forward Pro is designed to help traders achieve better and more robust results from their trading systems. It does this by assisting with the choice of settings that are most likely to produce an optimal walk forward optimization (WFO).

As an example, it specifically addresses the following questions:

- How many stages should be used in the WFO for my trading system?
- What ‘optimization to walk forward ratio’ should be used for the system?
- How many variables can be simultaneously optimized without over-optimizing (sometimes termed over-fitting) the system?
- How much price data is needed to achieve statistically significant results?

The complexity for the trader arises because the answers to each of these questions are interrelated and a decision for one of the questions will affect the optimal settings needed for the others.

Although walk forward optimization is widely considered to be one of the best techniques (if not the best technique) available to test and optimize trading systems, the effectiveness of the process is often reduced by choosing inappropriate settings for the process. This leaves the trader with a system that does not reach its full potential, or even worse, will result in an over-optimized system that has little or no chance of producing consistent profits in a real-money account.

This is where the Walk Forward Pro Machine Learning Module provides a solution.

WFP Machine Learning Algorithms

The balance of achieving the optimum settings for a WFO is determined by the following three factors:

- The ‘in-sample optimization effectiveness’
- The ‘out-of-sample walk forward confidence’
- The ‘adaptability score’

The relationship between these factors is an extremely complex one, and this can lead to confusion about how to achieve the right balance. It usually results in traders not using Walk Forward Analysis effectively, and not achieving the full potential from their trading systems. The complex relationships between these three factors also makes it difficult to determine the optimal balance using a rules based algorithm. Trade Like A Machine have striven to address these very real issues with the WFP Machine Learning Algorithms making use of AI (Artificial Intelligence). Following each test you undertake, ensure that you consider the ‘Machine Learning Review’ tab to identify what settings are likely to maximize the effectiveness of your walk forward analysis, and therefore help you to produce more robust systems.

Because the Machine Learning algorithms use Artificial Intelligence, they are able to adapt, learn and improve over time, meaning the more tests that are undertaken using them, the better they will become.
How does it work?
Before the Machine Learning Module can begin to suggest optimal settings for your trading system, it needs to review and analyse the results produced from your system during a ‘standard’ test (i.e. a test you have undertaken using manually selected settings).

Following the completion of every ‘standard’ test, the Machine Learning Algorithms will run automatically, and relevant options will be presented to the user on the ‘Machine Learning Review’ screen. This screen provides intelligence regarding the analysis of the test just undertaken, and highlights any potential issues (e.g. you will be notified if your settings have produced an ‘over-optimized’ system). It also provides suggested settings to overcome these issues.

The ‘Machine Learning Review’ Screen
The first section of the ‘Machine Learning Review’ screen shows a high level assessment of the test just undertaken. Specifically this shows how well the test scored against the following key metrics:

- Optimization Predictive Power
- Walk Forward Confidence
- Ability to Adapt

If traders are to produce robust systems that perform well in real money accounts, it is imperative that:

- There is sufficient ‘predictive power’ from the in-sample optimization phases. When there is a lack of statistical significance associated to these optimization phases, this reduces, or in extreme cases completely eliminates the predictive power of selecting the best parameter values. This then leads to poor performance in both out-of-sample walk forward tests, and also when traded in real money accounts. This is what the industry often calls 'over-optimization' or 'over-fitting', and is the number one reason why many traders fail to ever produce systems that perform well in live accounts. The ‘Optimization Predictive Power’ score shown on this screen provides an indication of whether the test settings that were used, provided sufficient ‘predictive power or not.
• The results achieved from the out-of-sample walk forward phases need to have a good level of 'Statistical Significance'. When there is a lack of statistical significance associated to the walk forward phases, this means the results of the cumulative walk forward tests cannot be trusted as being a true indication of the performance of your trading system. The ‘Walk Forward Confidence’ score shown on the screen provides an indication of whether the test settings that were used resulted in sufficient confidence or not.

• The WFO is conducted using settings that mean the trading system is able to adapt to changing market conditions. The ‘Adaptability score provides this information.

For a more comprehensive explanation of these 3 measures, see Appendix A.

Analysis of the WFO Test
The next section of the ‘Machine Learning Review’ screen provides a verbal, plain language explanation of these 3 metrics, and explains in detail what the implications are, based on the scores achieved.

The final section of the screen is where the results of the machine learning algorithms are presented to the user. Here, you will find suggestions of the settings that could produce significant improvements in your walk forward optimization.
Three different options are presented to the user to explore:

**Option 1:** Settings to achieve the ‘minimum acceptable balance’

In order to produce robust and profitable trading systems, it is imperative to ensure that the settings used for an optimization have good ‘predictive power’, and that the walk forward results can be ‘trusted’.

Option 1 identifies the settings required to produce the ‘absolute minimum’ level of statistical significance that is considered acceptable to achieve this. These settings will usually produce improvements in performance, especially if the current test suffered from over-optimization.

Option 1 therefore provides benchmark settings that are a good next step in the testing process. Following this test however, it is advisable that Options 2 and 3 are both also considered, since they will provide results that exceed the ‘minimum acceptable levels’ of statistical significance that option 1 provides. Options 2 and 3 will usually result in even more robust systems.

Option 1 uses the WFP Machine Learning Algorithms to achieve the optimal balance of the following:

- In-Sample Statistical Significance
- Out-of-Sample Statistical Significance
- Adaptability to changing market conditions

Specifically, the machine learning algorithms do the following:

1. Identify the maximum number of parameters that can be simultaneously optimized whilst still maintaining a ‘Minimum Acceptable’ statistical significance.
2. Choose the optimal value for the ‘number of stages’ based on the characteristics of the trading system.
3. Choose the optimal value for the ‘opt to walk forward ratio’ based on the characteristics of the trading system.
4. Note that the duration of the test is left unchanged from the previous test when using Option 1.
Option 2: Using an increased WFO Duration (more price data)

Option 2 attempts to build on the approach taken for Option 1, but illustrates how increases in statistical significance can be expected when increasing (doubling) the test duration by using additional price data.

If this additional price data is available, Option 2 will almost certainly produce more robust results than in Option 1.

Option 3: Reduced number of optimization variables

Attempting to optimize too many variables simultaneously is probably the most common reason for ‘over-optimization’ (sometimes called over-fitting). Many traders think that the more variables they optimize, the better the results will be. This is not the case - in fact the opposite is usually true, because optimizing too many variables simultaneously can severely impact the statistical significance of the in-sample optimization, which then results in an inability for the test process to select the best parameter values.

Option 3 builds on the settings in Option 1, but attempts to improve the performance of your system further by decreasing (halving) the number of variables being optimized.

Next, the WFP Machine Learning Algorithms are used to identify the optimal ‘number of stages’ and ‘opt to walk forward ratio’ to maximise the potential of the walk forward optimization process.

User intervention required when running the Machine Learning Options

Depending on the machine learning option chosen, the user will then be presented with either one or both of the following screens:

Parameter Selection Screen

When the chosen option requires a different number (usually a reduced number) of optimization parameters, compared to the original test, the user will need to choose the optimal parameters.
which parameters will be optimized in the next test, and which will not.

Many traders think the more variables that are simultaneously optimized, the better the results will be. This simply is not true (see side panel on ‘Optimization Best-Practice’).

**Important:** If you de-select a parameter so that it will no longer be optimized, **you must remember** to set the ‘default value’ field appropriately, since this is the value that will now be used for the next test. You can use the results from the previous test to determine what default value might be best.

![Optimization Parameters](image)

**Date Range Selection Screen**

Option 2 requires more data to be available for the test. When using this option, the user is presented with a choice of whether to use the most recent $n$ years or to specify a start and end date that complies with this duration of test. When using a specific start and end date, the end date defaults to the end date used in the previous test undertaken.

![Select Test Dates](image)

It is important when using this option, that you have sufficient price data for the extended duration, otherwise the settings suggested by the Machine Learning Algorithms will **not** be optimal.
**Machine Learning Tests**

When running a test based on settings suggested by the Machine Learning Algorithms, this is termed a ‘Machine Learning Test’ or ‘Machine Learning Optimization’ (MLO).

Following a MLO, an additional chart can be seen which is not available for a ‘Standard WFO’. This is the Machine Learning Comparison chart shown below, and it can be found on the ‘Out-of-Sample Walk Forward Results’ screen.

![Machine Learning Comparison Chart](image)

This chart shows a direct comparison between the out-of-sample equity charts of the previous ‘Standard WFO’ (the dotted line) and the new MLO (the solid line). This allows a direct visual comparison to be viewed allowing the user to determine if the machine learning settings have indeed improved the test results or not.

**Using the Machine Learning module on tests undertaken in the past**

The Machine Learning Algorithms have been designed not only to work on a test that has just been undertaken, but to also use the data from tests undertaken in the past, to provide intelligence to the user in exactly the same way. This is even the case for tests that were originally run before the Machine Learning algorithm was released into the product.

Whenever you open a past test from the ‘Previous Runs’ screen as shown below, the machine learning algorithms will be run immediately and the results shown on the ‘Machine Learning Review’ screen in the same way for a test that has just been run.
When re-opening a test that was undertaken in the past, you might notice that the proposed settings presented by the Machine Learning Algorithms are slightly different to results previously shown. This is because the Machine Learning Algorithms use AI (Artificial Intelligence), and so have the ability to learn and improve. If the algorithms have been ‘re-trained’ since you initially looked at the ‘Machine Learning Review’ screen for a particular test, then the results presented could be different (and hopefully improved).

IMPORTANT NOTES:

1. Both the Out-of-Sample and In-Sample Statistical Significance measures are calculated by Walk Forward Pro based on the assumption that the trades produced during the testing are independent and un-correlated i.e. there are not multiple trades that are almost identical in nature (with similar entry times and price levels, and/or similar exit times and price levels). If your trading system does produce many trades that are almost identical (e.g. by scaling into positions with many similar micro trades, and/or scaling out of trades in the same way), then the statistical significance scores shown will be inflated and cannot be relied upon. The ‘real’ statistical significance values will in reality be less (significantly less, if there are many trades that are similar).

2. The machine learning algorithms are deployed as part of Trade Like A Machine’s cloud services, and are not built directly into the installed client product. This allows the algorithms to learn, adapt, and improve over time, based on their use by all Walk Forward Pro customers. This does mean however, that there is a requirement to have internet access at the time they are used (which is at the end of a test or when opening a previous test). If you wish to use the intelligence from the machine learning algorithms, please ensure that your Wi-Fi is turned on, and that any firewalls you use are configured to allow internet access to Walk Forward Pro.
Appendix A

**OPTIMIZATION PREDICTIVE POWER (IN-SAMPLE STATISTICAL SIGNIFICANCE)**

If traders are to produce robust systems that perform well in real money accounts, it is imperative that there is sufficient predictive power from the in-sample optimization phases. To ensure this is the case, the optimizations must be undertaken based on 'statistically significant' data.

**Why is it so important?**

When there is a lack of statistical significance associated to the optimization phases, this reduces, or in extreme cases completely eliminates the predictive power of selecting the best parameter values. When this is the case, the selection of parameter values tends to be made based more on randomness and chance, than on the effectiveness of the actual parameter values chosen. This then leads to poor performance in both out-of-sample walk forward tests, and also if traded in real money accounts. This is what the industry often calls 'over-optimization' or 'over-fitting', and is the number one reason why many traders fail to produce systems that perform well in live accounts.

An ‘Optimization Predictive Power Score’ of below 1 means it is likely that the parameter values were chosen based more on randomness and chance than the effectiveness of the actual parameter values. The trader should attempt to get this score as high as possible, in order to improve the effectiveness of the optimization phases.

**Improving the Optimization Predictive Power**

In order to improve the ‘Optimization Predictive Power’ the user could do one or more of the following:

1. The most effective way is to reduce the number of parameters being simultaneously optimized. Many traders feel that this will decrease the ability to find the 'optimum' trading system. Actually, the opposite is true. Better out-of-sample results will usually be produced when the number of parameters being optimized is low.

2. Increase the overall walk forward analysis duration (e.g. instead of testing over just 5 years, obtain data so that you can test over 10 years).

3. Increasing the 'Optimization to Back Test Ratio' will improve this score. (The resulting shorter Walk Forward periods will however have a detrimental effect on the 'Walk Forward Confidence' score and so this is somewhat of a balancing act). The Walk Forward Pro Machine Learning Algorithms have been specifically developed to find the optimum balance for you, and to suggest an 'Optimization to Back Test Ratio' that is likely to achieve the best balance and improve the overall test results.

4. Decreasing the number of stages in your walk forward analysis will also have a beneficial effect on this score, but again, this will have a detrimental effect on the 'Walk Forward Confidence' and is again therefore a balancing act. Once again, the Walk Forward Pro Machine Learning Algorithms have been built to perform this 'balancing' task, and suggest the number of stages that is likely to deliver the optimum overall results.
### Overview

If traders are to produce robust systems that they can trust, it is imperative that the results achieved from the out-of-sample walk forward phases have 'Statistical Significance'.

### Why is it so important?

When there is a lack of statistical significance associated to the walk forward phases, this means the results of the cumulative walk forward tests cannot be trusted as being a true indication of the performance of your trading system.

Note that the 'Walk Forward Confidence Score' is calculated and solely based on analysis of the walk forward phases, and so gives a level of confidence in the results displayed on the ‘Out-of-Sample Walk Forward’ screen. This score does not reflect the significance of the in-sample optimizations or the parameter selection robustness (for this, see the 'Optimization Predictive Power Score').

A 'Walk Forward Confidence Score' below 1 means that the results achieved in the out-of-sample walk forward test will almost certainly not be representative of the results that would be experienced when trading the system in a live account. The trader should attempt to get this score as high as possible, in order to have confidence in the results.

### Improving the Walk Forward Confidence

In order to improve the ‘Walk Forward Confidence’ the user could do one or more of the following:

1. Increase the overall walk forward analysis duration (e.g. instead of testing just 5 years, obtain data so that you can test over 10 years). This is the most effective way of improving this score.

2. Decreasing the 'Optimization to Back Test Ratio' will improve this score. (The resulting shorter Optimization periods will however have a detrimental effect on the 'Optimization Predictive Power' and so this is somewhat of a balancing act). The Walk Forward Pro Machine Learning Algorithms have been developed to find the optimum balance for you and to suggest an 'Optimization to Back Test Ratio' that is likely to improve test results.

3. Increasing the number of stages in your walk forward analysis will also have a beneficial effect on the out-of-sample statistical significance, but again, this will have a detrimental effect on the 'Optimization Predictive Power' and is again therefore a balancing act. Once more, the Walk Forward Pro Machine Learning Algorithms have been built to perform this ‘balancing’ task, and suggest the number of stages that is likely to deliver the optimum overall results.
Overview

One of the reasons that walk forward analysis can be so effective is that it allows the optimizations and associated walk forward phases to be based on recent market conditions. It allows systems to adapt to changing markets by selecting parameters matched to those recent conditions.

How does it work?

The Adaptability Score provides a measure of the degree to which your walk forward analysis settings take advantage of this capability i.e. it measures how well your test settings prove that your system is capable (or otherwise) of adapting when market conditions change.

The Adaptability Score is calculated by determining how each stage of the walk forward analysis takes advantage of ‘new’ data, not used in previous stages, and therefore the exposure of each stage to different market conditions.