**Trading Instrument:** ES (S&P Emini futures contract)

**Tick Increment:** .25 (1/4 of 1 point) = $12.50/tick ($50/point)

**Objective:**

Analyze the exported block data to find a profitable predictable edge to implement into my trading algorithm.

**How Will We Know That the Objective Has Been Achieved?**

The analytics model will have to produce a predictability accuracy % that results in a profitable outcome. For that to happen the model must consider the predictability accuracy %, the specific number of ticks of market movement accurately predicted prior to a certain amount of market movement in the opposite direction. Another way to state this is, the prediction accuracy percentage and total gain predicted by the model must be greater than the amount of loss incurred when the prediction is inaccurate.

In addition to the above, the model must be able to be provided to my programmer to be implemented into my strategy. The results must be reproducible and implementable. Meaning it cannot be a black box or locked code, the code would have to be provided in its entirety so that I can use it.

**Targets and Assumptions**

Target Predicted Profit Amount: 8 ticks (2 points) or Greater from the Close price of the Breakout Bar (definitions are provided below).

A failed prediction (loss) would be if the market moved 12 ticks in the opposite direction from the Close price of the Breakout Bar prior to achieving the 8 tick target.

NOTE: If through optimization of the model it can predict a winning scenario close to what is described below, the model and algorithm may still be feasible.

**Data Format**

The Block Data will record data on each bar that is within or the breakout from a block. This shall contain the following fields:

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| block\_id | Identifies which block the bar belongs to, currently in the form of the time for the first bar. |
| bar\_time | The date and time corresponding to the closing of the bar. |
| bar\_number | Identifies where the bar falls within the block. 0 identifies the first bar, with the id incremented for each successive bar. The breakout bar is identified by -1, because it is important, but should often be considered on its own, and sometimes the block should be considered excluding it. |
| open | The open price of the bar. |
| high | The high price of the bar. |
| low | The low price of the bar. |
| close | The close price of the bar. |
| price[0-6] | Specifies the prices levels at which contract were bought and/or sold within the bar. A value of 0 is used for indices without data. |
| buy[0-6] | The volume of contracts “bought” at the specified price level. The indices correspond to the price columns. |
| sell[0-6] | The volume of contracts “sold” at the specified price. The indices correspond to the price columns. |
| type | An alphanumeric identifier corresponding to how the block was created. |
| upper\_dist | The distance in price to the next highest point of interest. |
| lower\_dist | The distance in price to the next lowest point of interest. |

Additionally, tick data will be provided over the same time period for the evaluation of results. This data shall be in the following format:

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| Date | The date at which the order was executed. |
| Time | the time at which the order was executed. |
| Price | The price at which the order was executed. |
| Volume | The number of contracts in the transaction. |

**Defining Terms**

These terms are used within the calculation section. An in-context definition is provided in the table below for clarity.

|  |  |
| --- | --- |
| **Term** | **Context and Definition** |
| Bar | Each row in the block data set represents one bar of data. This name will be used interchangeably with row in the calculation table. |
| Block | The collection of bars (rows) with the same *block\_id,* excluding the bar with *bar\_number* = -1. |
| Breakout Bar | A bar with *bar\_number* = -1. This is the first bar to exceed the price bounds of the block (identified by the *block\_id*). |
| Up Close | A bar that has the close value greater than the open value |
| Down Close | A bar that has the open value greater than the close value. |

**Calculations for Desired Features and Intermediate Values**

Instructions for calculating our desired features and intermediate values used in the calculations.

|  |  |  |  |
| --- | --- | --- | --- |
| **Type** | **Name** | **Feature Type** | **Calculation** |
| Integer | Bar Buy Volume | Intermediate Value | Sum the values of all *buy* columns within the bar (row). |
| Integer | Bar Sell Volume | Intermediate Value | Sum the values of all *sell* columns within the bar (row). |
| Integer | Bar Total Volume | Intermediate Value | Sum of *Bar Buy Volume* and *Bar Sell Volume*. |
| Integer | Block Buy Volume | Feature | Sum the *Bar Buy Volume* columns for all rows that share a *block\_id*, excluding the row with *bar\_number* = -1. |
| Integer | Block Sell Volume | Feature | Sum the *Bar Sell Volume* columns for all rows that share a *block\_id*, excluding the row with *bar\_number* = -1. |
| Integer | Total Block Volume | Intermediate Value | Sum the *Block Buy Volume* and *Block Sell Volume*. |
| \*Diagram Below | Block Volume Profile | Intermediate Value | A set of key-value-value triplets, of arbitrary size. The key is the price, the first value is the buy volume, the second value is the sell volume. For each value in the *price* columns across all bars in the block, sum the corresponding *buy* and *sell* columns. |
| Integer | Buy Volume at High Volume Price Level in Block | Feature | Using the Block Volume Profile, determine the price level with the highest total volume. Equal to the *buy volume* at this price level. |
| Integer | Sell Volume at High Volume Price Level in Block | Feature | Using the Block Volume Profile, determine the price level with the highest total volume. Equal to the *sell volume* at this price level. |
| Integer | Volume Difference at High Volume Price Level in Block | Feature | Using the Block Volume Profile, determine the price level with the highest total volume. Equal to the *buy volume – sell volume* at this price level. |
| Enum | Close Direction of Highest Volume Candle | Feature | For each block, identify the bar with the greatest *Bar Total Volume*. For this bar, determine if it has an up close or a down close. |
| Enum | Close Direction of 2nd Highest Volume Candle | Feature | For each block, identify the bar with the 2nd greatest *Bar Total Volume*. For this bar, determine if it has an up close or a down close. This value should not be used for blocks containing only one bar. |
| Boolean | Divergence of Highest Volume Candle | Feature | For each block, identify the bar with the greatest *Bar Total Volume*. For this bar, determine if the candle is divergent. The candle will be considered divergent if:   * The *Close Direction of Highest Volume Candle* is and Up Close AND *Bar Sell Volume > Bar Buy Volume*, OR * The *Close Direction of Highest Volume Candle* is and Down Close AND *Bar Buy Volume > Bar Sell Volume* |
| Boolean | Divergence of 2nd Highest Volume Candle | Feature | For each block, identify the bar with the 2nd greatest *Bar Total Volume*. For this bar, determine if the candle is divergent. The candle will be considered divergent if:   * The *Close Direction of Highest Volume Candle* is and Up Close AND *Bar Sell Volume > Bar Buy Volume*, OR * The *Close Direction of Highest Volume Candle* is and Down Close AND *Bar Buy Volume > Bar Sell Volume*   This value should not be used for blocks containing only one bar. |
| Integer | Total Volume of Breakout Bar | Feature | For each block, is equal to the *Bar Total Volume* for the breakout bar (*bar\_number* = -1). |

**Variation** – we would like to also run the model using ALL the same features listed above, however, with one variance, we would like to include the Breakout Bar as part of the block.

We would like to plot the prediction accuracy of each feature against the *Total Block Volume* to determine if it affects predictability.

\*Diagram showing the calculation of block volume profile for a block with two bars.

