Manual: 4.2. Correction and Validation Rules

When the machine learning algorithm learns the model, it retrieves the historical data from the database point-by-point. Each point is first <u>corrected</u> and then <u>validated</u>. Only valid points are seen by the learning method and invalid points are ignored.

Think of a valve and the tag that measures how open that valve is. A fully open valve records 100% openness and a fully closed valve records 0% openness. As a result, the minimum and maximum values for this tag are 0 and 100 respectively. However, we find in practice that due to a variety of reasons we do get numbers stored in the database that are a little below 0 or a little above 100. These are not really measurement errors or faulty points. We should not throw this data away but correct this data to what we know it to be. A measurement below 0 is actually 0 and a measurement above 100 is actually 100. We know this by the nature of the thing being measured and so we use a correction rule. <u>Correction rules</u> are rules that change the numerical value of a tag to a fixed value, if the value recorded is above or below some threshold. By default, there are no corrections applied to the data. If corrections are necessary, you must specify them manually.

The machine learning algorithm should only learn operational conditions that are reasonable in the sense that it would be ok if the model were to decide to bring the plant to that condition in the future. Any exceptional conditions that the plant was actually in at some point but should not really get into again, ought to be precluded from learning. Also any condition in which the plant is offline, not producing, in maintenance or any other non-normal condition ought to be marked invalid. A <u>validity rule</u> requires that a certain tag be within specified limit values. By default, every tag must have a value between its minimum and maximum for the point to be valid.

In addition to the default validation rules, we may manually add others. These additional rules do not have to always apply and so we can add a specification that this rule only applies when some tag is above or below a certain value. This applicability feature allows you to enter more complex rules depending on conditions. For example you could formulate the requirement that the process have a high temperature at full-load and a low temperature at half-load.