Manual: 8.2. Tag

A sensor is a physical device that converts some aspect of a physical situation into a numerical value relative to some agreed-upon scale. A temperature sensor, for example, will output the temperature of its ambient medium as a numerical value on the Celcius or Fahrenheit scale.

It is important to note that sensors suffer from several sources of error that must be taken into account when interpreting the numerical value in question.

First, each sensor has a measurement uncertainty, specified by its manufacturer, that illustrates that the same physical situation may lead to results anywhere within a numerical range of this size.

Second, every sensor suffers from drift over time. With age, accumulation of dirt and other factors, the numerical output value has a tendency to steadily rise or fall. This leads to the need to periodically calibrate sensors or replace them.

Third, a sensor may have a systematic error due to its placement. For example, a temperature sensor placed on the outside of a large vessel will measure the temperature there but this temperature will not be equal to the average temperature in the vessel. For machine learning analysis, this effect usually does not influence model quality as the error is always constant. However, this error is significant if a physical interpretation is needed. For example, if we know that the best temperature is 100 degrees and we measure 98 degrees, we may be tempted to inject heat but this may not be necessary as the outside of the vessel may legitimately be 2 degrees cooler than its average internal temperature.

A measurement is an individual numerical result delivered by a particular sensor. A measurement is thus always three pieces of information: The time at which it was taken, the tag it is relevant to, and the value that was recorded.

A tag encapsulates a regular measurement made by a sensor and all the data associated to that operation. If the sensor must be exchanged or recalibrated, the tag remains the same. All the data ever recorded relative to that quantity is stored under the tag's identification. Tags are therefore the currency on which all process models are built.

A sibling to a tag is another tag that measures the same quantity in the same manner in roughly the same location. We desire to have siblings in the case of important measurements for which we cannot afford a sensor malfunction or failure. A normal industry standard is to install three sensors and, at any one time, use the average of the two measurements that most closely agree as the true value measured by the triplet of sensors. In case the siblings disagree significantly with each other, we have two potential problems. Either a sensor is malfunctioning or the equipment truly exhibits this deviation in between the (usually very close by) locations of the sensors in which case the equipment is probably in an unhealthy state.