Data Collection and Modeling

Course 3: Data Acquisition

Definition

“Data acquisition is the process of sampling signals that measure real world physical conditions and converting the resulting samples into digital numeric values that can be manipulated by a computer. Data acquisition systems (abbreviated with the acronym DAS or DAQ) typically convert analog waveforms into digital values for processing.” (Dataforth)

“Data acquisition (DAQ) is the process of measuring an electrical or physical phenomenon, such as voltage, current, temperature, pressure, or sound.” (National Instruments)

Other Definitions

- The process that uses a data acquisition system (DAS) to measure and digitize real world measurements/signals data
- The process of collecting data that has already been measured and digitally transformed

Depending on the project both these processes can be involved together with synthetic data generation (synthetic data obtained by simulation and generation to have enough instances for training a ML model)

Related Terms

- **Measurement** - quantitative determination of a physical characteristic. Usually it is the conversion of a physical quantity or observation to a domain where the value can be determined (by human or computer).
- **Instrumentation** – devices for converting a physical quantity to a quantity observable by a human or computer.
- **Data Acquisition** - Gathering information from measurement sources, such as sensors/transducers.
- **Sensor** - device that converts physical phenomena such as temperature, vibration, and acoustics to electrical signals.

Applications of DAQ

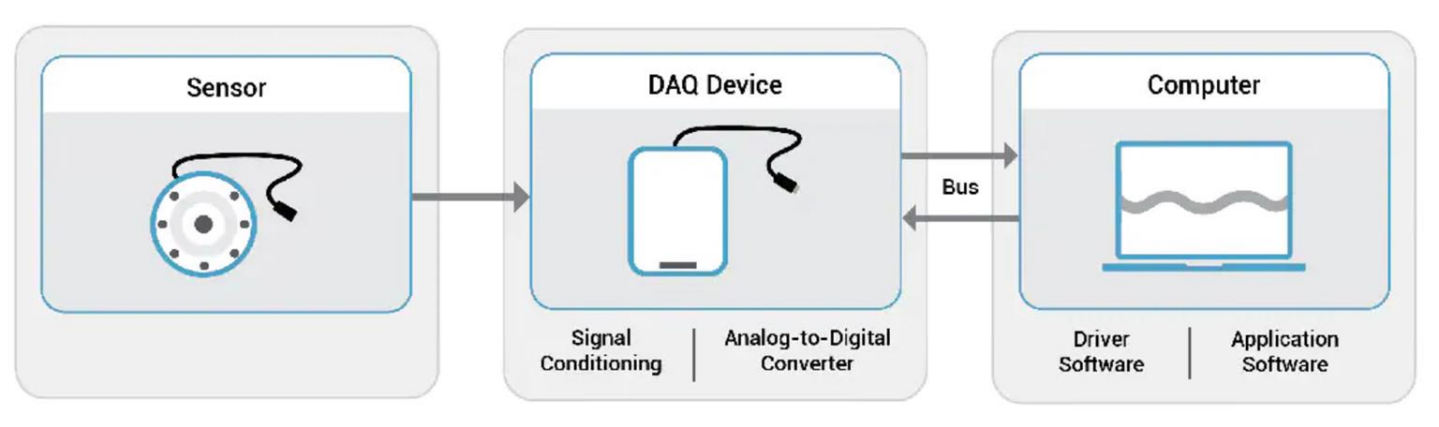
- **Scientific Exploration**: collect information for experiments, simulations, observations
- **Industrial Efficiency**: monitoring manufacturing processes for quality assurance, optimizing overall efficiency
- **Environmental Insights**: tracking critical factors: air quality, water levels, soil conditions; allows effective environmental management and timely disaster prediction
- **Healthcare and Biomedical Studies**: monitor vital signs, acquire physiological data, help in diagnostic accuracy
- **Automotive Evaluation**: testing vehicle performance, safety features, and efficiency across diverse scenarios

Challenges of DAQ

- **Accuracy and Calibration**: sensors and instruments must be accurate and well-calibrated
- **Data Integrity**: avoid signal interference and noise for data integrity
- **Sampling Rates**: crucial to choose it not to miss important data but also not to overload the system
- **Compatibility and Integration**: compatibility issues when integrating sensors and instruments from different manufacturers

DAQ System

consists of sensors, DAQ measurement hardware, and a computer with programmable software



Source: Omega

Sensors

Sometimes called transducers

“a transducer is a device that converts a signal from one physical form to a corresponding signal that has a different physical form. Therefore it is an energy converter” (Sensors and Signal Conditioning)

Sensors or transducers interact with the subject measured, either directly or indirectly (contact or non-contact). They convert the physical values to electric signals. The type of sensors used in a DAQ system varies based on the nature of its application.

Sensors

- **Active sensors**: produce output by means of external excitation supply; the physical properties changes depend on applied external effect
- **Passive sensors**: produce output without external excitation; they don’t need external voltage/power supply

- **Analog sensors**: produce analog output as continuous signal with respect to time
- **Digital sensors**: produce discrete non-continuous with respect to time; output signals converted and transmitted digitally

Sensor Types

Phenomena	Sensors
Force, pressure	Load cells Strain gauges
Fluid flow	Rotational flowmeters Ultrasonic flowmeters
Light	Photoconductive cells Vacuum tube photosensors
Position	Linear voltage differential transformers Optical encoders Potentiometers
Sound	Microphone
Temperature	Thermocouples Thermistors Resistance temperature detectors

source: electricaltechnology.org

Signal Conditioning

The process of optimizing the signal

Sometimes the signal can contain noise, can be only partially measured (weak)

- Calibration
- Linearization/Offsetting
- (External) Excitation
- Amplification
- Filtering

DAQ Hardware

The device connected between sensor and computer; can be connected to computer on various ports/slots (USB, PCI-Express)

Receives (analog) output signals from sensors and transforms/converts them to digital signals that can be processed by computers

Choosing the Right DAQ Hardware

1. What types of signals are used?
2. Is signal conditioning needed?
3. What is the sampling rate?
4. What resolution is needed?
5. What is the accuracy?

What types of signals are used?

Different types of signals have to be measured/generated in different ways; sensors convert physical phenomena into measurable electrical signals (voltage or current). Sensors can also receive a measurable electrical signal to produce physical phenomena.

Functions of DAQ Devices:

- Analog inputs measure analog signals
- Analog outputs generate analog signals
- Digital inputs/outputs measure and generate digital signals
- Counters/timers count digital events or generate digital pulses/signals

There are both dedicated and multifunctional devices that can be used

ADC

Takes environment (analog) data from sensors and converts it to a digital form (discrete levels) usable in computer environment.

Characterized by the number of bits (4, 8, 12, 16, …). The larger this value, the greater number of discrete digital values can be represented (resolution).

- Successive Approximation (SAR)
- Delta-Sigma
- Dual Slope
- Pipelined
- Flash

Is signal conditioning needed?

General-purpose DAQ devices usually measure or generate voltages in the ranges of +/-5 V , +/-10 V.

There are sensors that generate signals that are difficult to measure directly with general purpose DAQ devices; these sensors require signal conditioning: amplification, filtering, linearization.

	Amplification	Attenuation	Isolation	Filtering	Excitation	Linearization
Thermocouple	x			x		x
Thermistor	x			x	x	x
RTD	x			x	x	x
Strain Gage	x			x	x	x
Load, Pressure, Torque (mV/V, 4-20mA)	x			x	x	x
	x			x	x	x
Accelerometer	x			x	x	x
Microphone	x			x	x	x
Proximity Probe	x			x	x	x
LVDT/RVDT	x			x	x	x
High Voltage		x	x			

source : ni.com

What is the sampling rate?

Sampling rate represents the speed at which the ADC takes samples of a signal.

Each application's sampling rate depends on the maximum frequency component of the signal that needs to be measured or generated.

The Nyquist Theorem stipulates that, in order to accurately rebuild the signal, twice the highest component frequency is needed.

The practical experiments show that, to obtain the correct shape of the signal, the sample rate should be at least 10 times the maximum frequency.

What resolution is needed?

The resolution represents the smallest detectable change in the signal. It represents the number of binary levels an ADC can use to represent a signal.

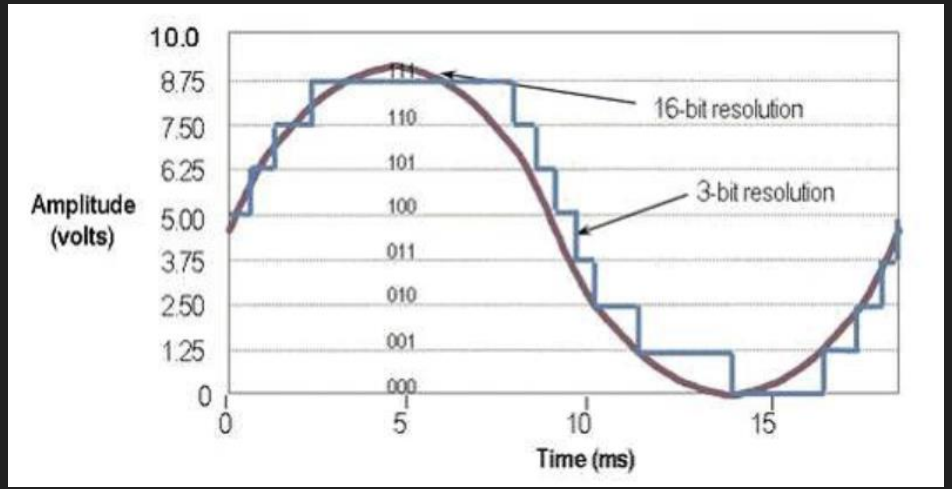
Imagine a 3-bit ADC and a 16-bit ADC:

- 3-bit ADC can represent

  8 (2^3) discrete voltage levels

- 16-bit ADC can represent

  65,536 (2^{16}) discrete levels



source: [ni.com]

What is the accuracy?

Accuracy can be stated as the uncertainty (amount) in a measurement with respect to an absolute standard.

Accuracy specifications usually contain the effect of errors due to gain and offset parameters. Offset errors can be given as a unit of measurement such as volts or ohms and are independent of the magnitude of the input signal being measured.

Gain errors depend on the magnitude of the input signal and are expressed as a percentage of the reading.

Total accuracy is equal to the sum of the two(example ±(0.1% of input +1.0 mV))

DAQ Features

- Digitization
- Multiplexing
- Conversion
- Transmission
- Binary output

Data Acquisition Options

- **Data Loggers**: data recording over a time period; it a self contained DAQ system with an embedded processor and software. Stand-alone, portable devices
- **Data Acquisition Devices**: can be connected on various ports (USB, Ethernet, PCI); contains signal conditioning and ADC but the software is on a connected computer
- **Data Acquisition Systems**: used for complex systems; integrates multiple sensors (different types) and synchronizes them.

Remote Data Acquisition

- Gather data from remote location and send it to a central machine for storage and processing
- Composed of sensors, DAQ hardware, communication hardware
- Allows monitoring and controlling equipment from a distance
- Offers
  - real-time data collection helping decision about remote equipment operation and maintenance/service
  - Monitor conditions in environments where a human operator can’t have access/can’t work securely

Remote Data Acquisition

- Web-based: accessible using a browser; easy to use
- Desktop based: installed on a desktop computer (RDP/SSH)
- Embedded: integrated with other devices (industrial controllers)
- Wireless: collected data from sensors is sent wirelessly to a controller; suitable for non wireable environments

Data can be collected simultaneously from several locations, saving time and money; Security can be a concern, especially with sensitive data; Networking plays a major role (redundant communication channels?)

Can be used for monitoring patients at home, power plants, collecting scientific data

Data Acquisition Methods for ML

Data can be collected to be cleaned and processed as training data for ML. We can identify 3 methos of data sourcing that can be used in any combination:

- Manual collection
- Data Procurement
- Open Source Warehouses