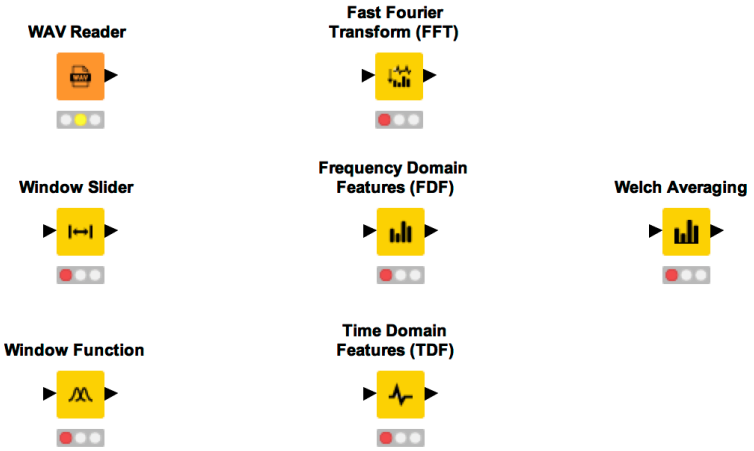
Signal Processing Nodes

Introduction and Software Installation

Signal Processing Nodes for the KNIME Analytics Platform donated by AI.Associates

ai. associates



AdvancedAnalytics .Academy

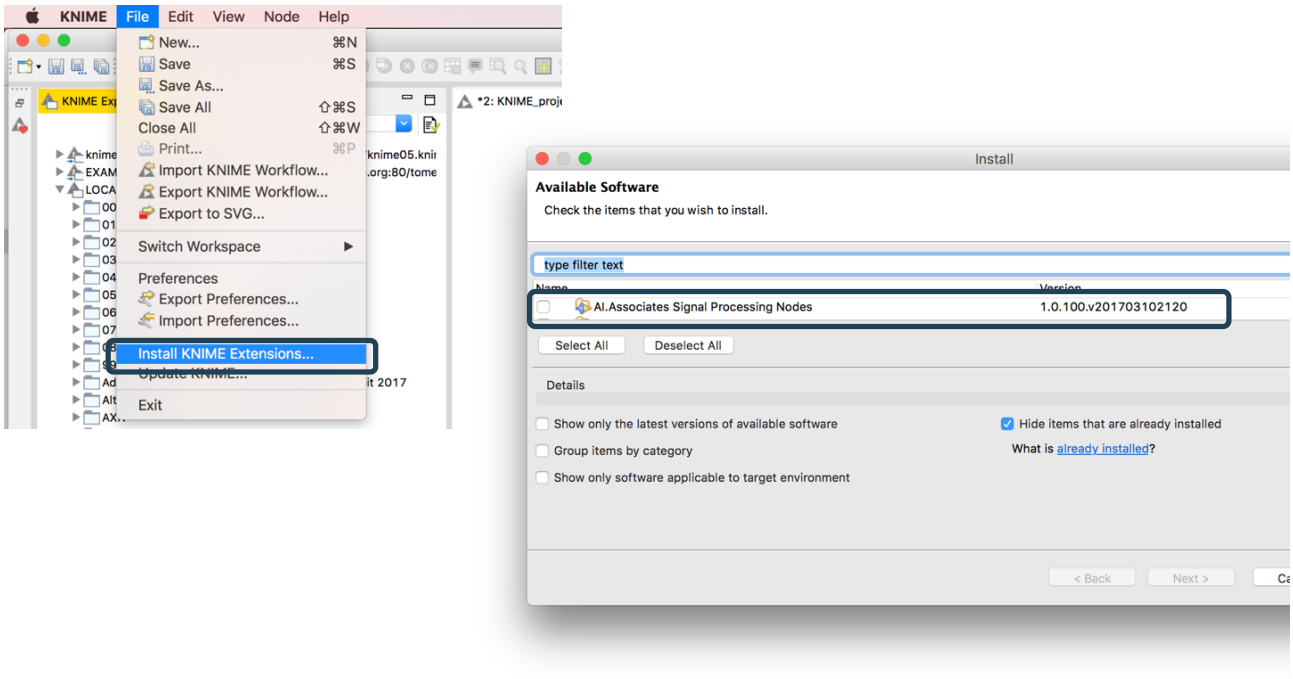
Signal Processing has never been that easy

With KNIME Signal Processing nodes you can easily

- Acquire, measure, and analyze signals from many sources, like audio, smart sensors, instrumentation, and IoT devices.
- Combine digital signal processing techniques with machine learning algorithms.
- Provide instant insights into signals without writing a single line of code.

AdvancedAnalytics .Academy

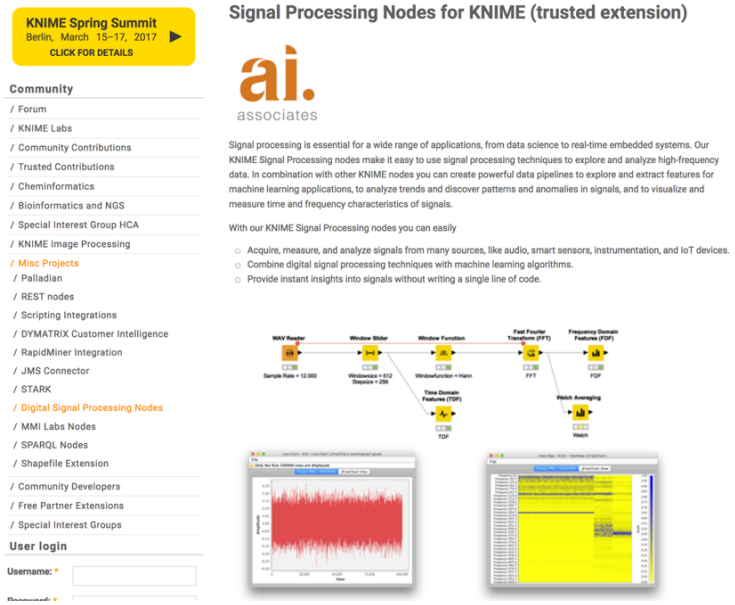
Installation of Signal Processing Nodes



AdvancedAnalytics .Academy

More about AI.Associates Signal Processing Nodes

https://tech.knime.org/community/digital-signal-processing-nodes

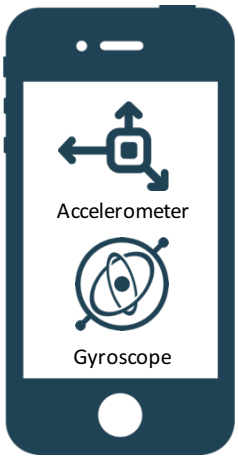


AdvancedAnalytics .Academy

IoT Analytics Use Case Activity Detection with DSP & Machine Learning

IoT Analytics

Overview



Smart Phone

Digital Signal Processing & Machine Learning

Activity Classification

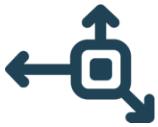
Experimental Setup*

- The experiments have been carried out with a group of 30 volunteers within an age bracket of 19-48 years.
- Each person performed six activities (WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING) wearing a smartphone on the waist.
- The embedded accelerometer and gyroscope capture 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50 Hz.
- The experiments have been video-recorded to label the data manually.
- The obtained dataset has been randomly partitioned into two sets, where 70% of the volunteers were selected for generating the training data and 30% for generating the test data.

* based on UCI HAR Dataset (https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones)

Accelerometer Sensor

- Accelerometer sensors measure linear acceleration along three perpendicular axes, X, Y, and Z.
- All smart phones are equipped with a modern accelerometer that can measure acceleration in three perpendicular axes.



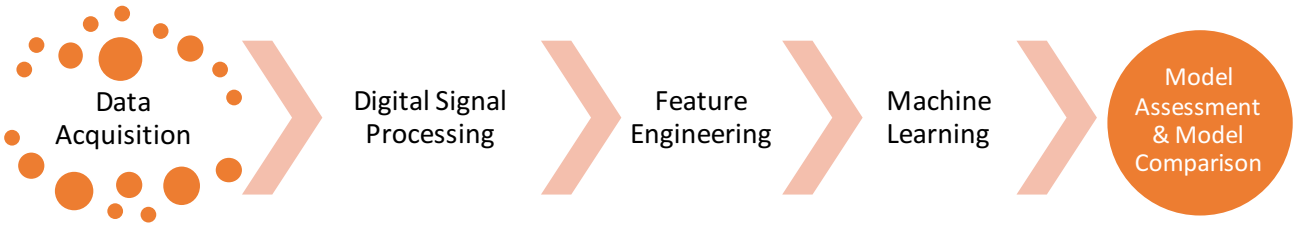
- The accelerometer reports acceleration in units of meter per second squared.
- Note that the measurement also includes the Earth's gravity (when the smart phone is placed on a flat table, it should display about 9,807 m/s^2 along the Z axis, and 0 along the X and Y axes).

Gyroscope Sensor

- Gyroscope sensors measure rotational velocity along the Pitch (X-axes/left-right), Roll (Y-axes/bottom-up) and Yaw (Z-axes/perpendicular to the face of the device) axes.
- It depends on the property of rotating mass as illustrated in the following schematic drawing of the classical mechanical gyroscope

- The gyroscope reports angular velocity in units of radians per second.

Analytical Process

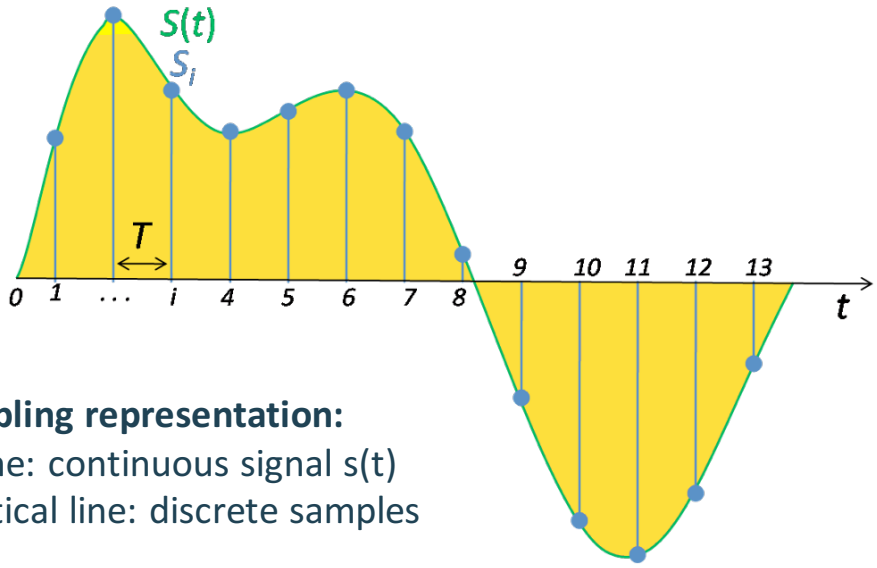


1.
Data Acquisition

1. Sensor Data Acquisition

- Beyond accessing **sensor data** in signal processing it is essential to know the **sample rate**.
- Sampling is the reduction of a continuous analog signal to a discrete signal (an example is the conversion of a sound wave (a continuous signal) to a sequence of samples (a discrete-time signal)).
- A sample is a value or set of values at a point in time and/or space.
- The **sample rate f_s** is the average number of samples obtained in one second (*samples per second*), thus **f_s = 1/T**.

1. Sensor Data Acquisition – Sample Rate



Signal sampling representation:

- Green line: continuous signal s(t)
- Blue vertical line: discrete samples

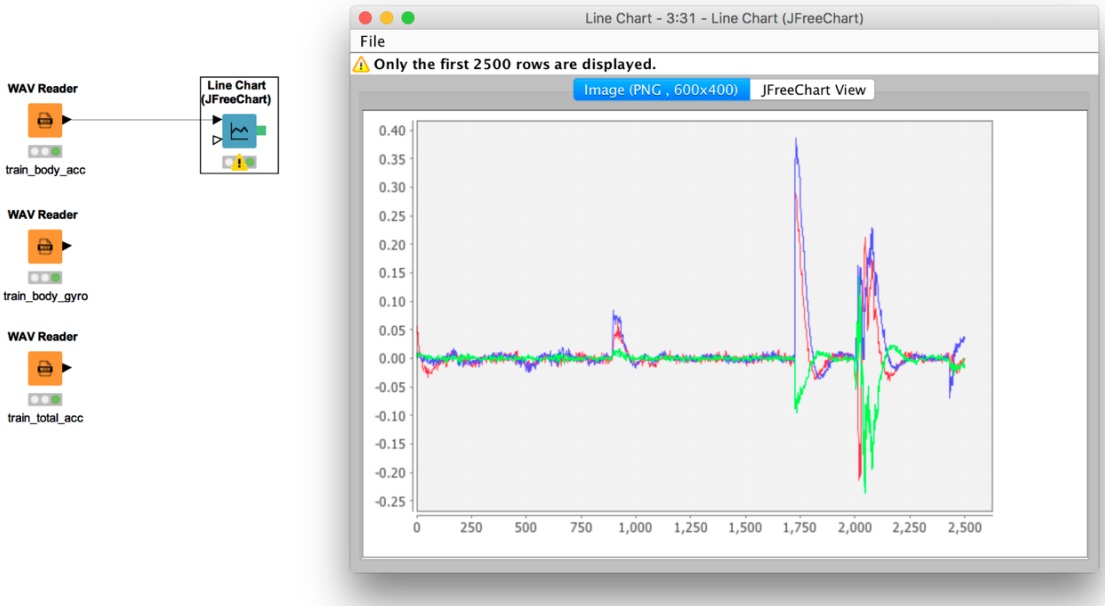
1. Sensor Data Acquisition – Sample Rate

- Most sampled signals are not simply stored and reconstructed.
- The fidelity of a theoretical reconstruction is a customary measure of the effectiveness of sampling.
- That fidelity is reduced when $s(t)$ contains frequency components whose periodicity is **smaller than 2 samples**; or equivalently the ratio of cycles to samples exceeds ½.
- The quantity $\mathbf{f_s/2}$ is known as the **Nyquist frequency** of the sampler.

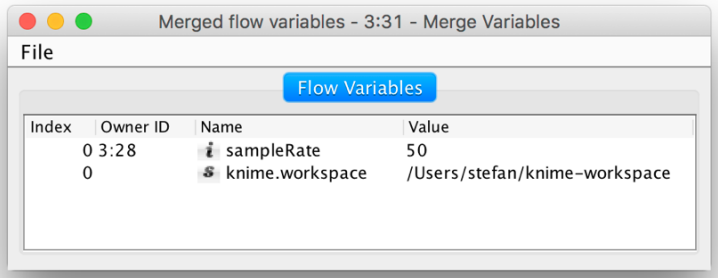
1. Sensor Data Acquisition – WAV Reader Node

- A Wave file is a standard PC audio file format (created by Microsoft) for everything from system and game sounds to CD-quality audio.
- A Wave file is identified by a file name extension of WAV (.wav).
- In addition to the uncompressed raw audio data, the Wave file format stores information about the file's number of channels (e.g. mono, stereo or more), sample rate, and bit depth.
- The KNIME WAV Reader node (donated by AI.Associates) provides both access to the raw sensor data of each channel and the sample rate.

1. Data Acquisition & Exploration of sensor data



1. Data Acquisition – Sample Rate (provided as a Flow Variable)



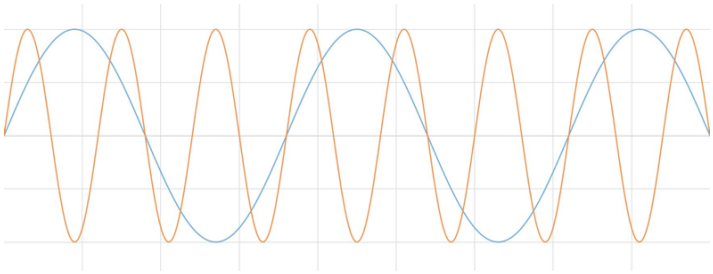
- The **sample rate f_s** is the average number of samples obtained in one second (*samples per second*), thus **$f_s = 1/T$**.
- If f_s = 50 Hz then Nyquist frequency is 25 Hz.

2. Digital Signal Processing

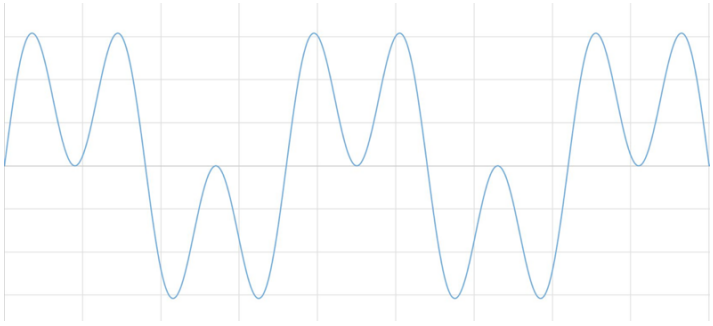
2. Time Domain, Frequency Domain, and FFT

- The Fourier transform can be powerful in understanding everyday signals and troubleshooting errors in signals.
- Although the Fourier transform is a complicated mathematical function, it isn't a complicated concept to understand and relate to your measured signals.
- Essentially, it takes a signal and breaks it down into sine waves of different amplitudes and frequencies.

2. All signals are the sum of sines

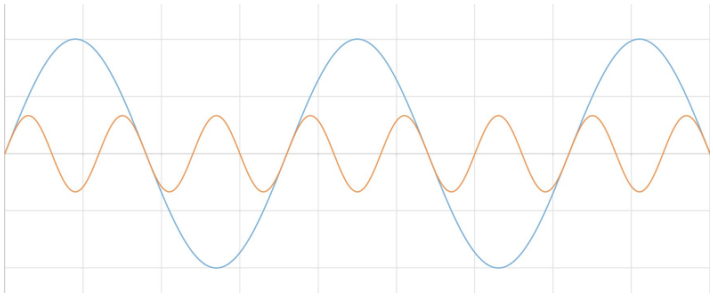


2 separated sine waves, where one is three times as fast as the other, or the frequency is 1/3 the first signal.

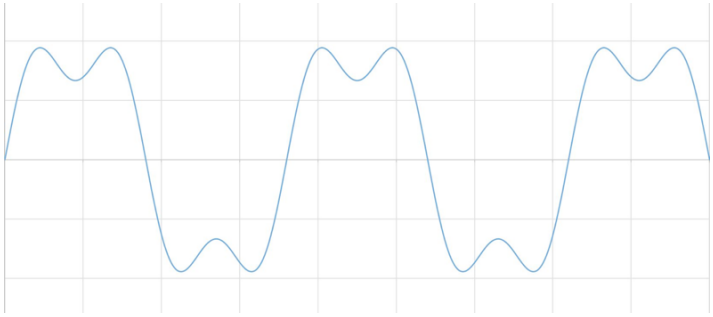


2 added sine waves, we receive a different signal.

2. All signals are the sum of sines



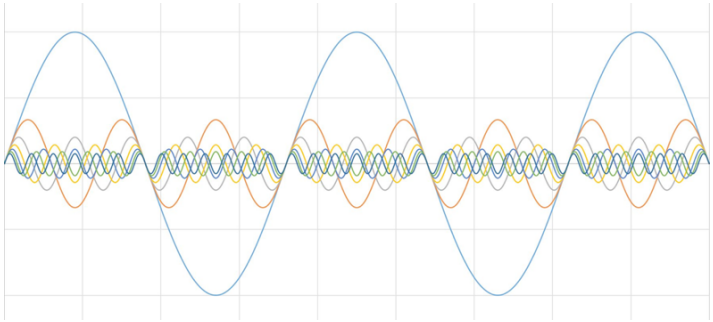
2 separated sine waves, where the frequency is 1/3 and amplitude 1/3 the first signal.



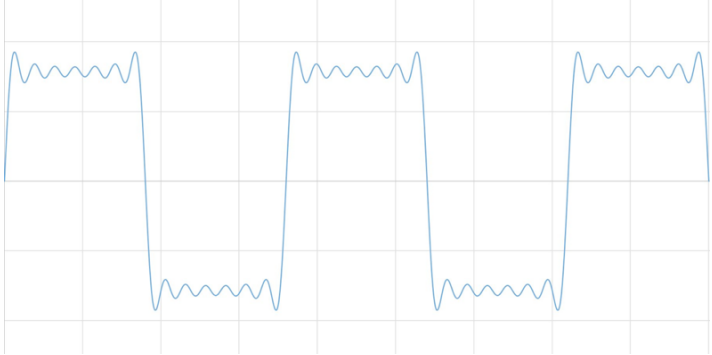
Adjusting the amplitude when adding signals affects the peaks.

AdvancedAnalytics .Academy

2. All signals are the sum of sines



Adding a third signal that was 1/5 the amplitude and frequency of the original signal…

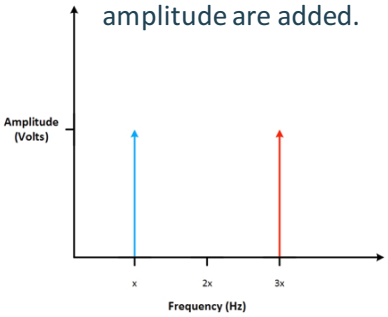


All signals in the time domain can be represented by a series of sines (square wave as a sum of sines)

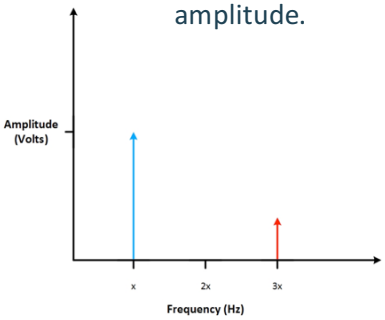
AdvancedAnalytics .Academy

2. FFT – Frequency Domain Representation

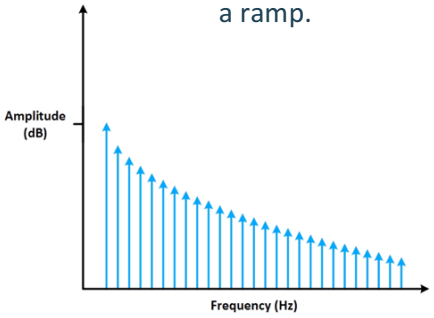
Two sine waves of equal amplitude are added.



The highest spike is the frequency of the largest amplitude.



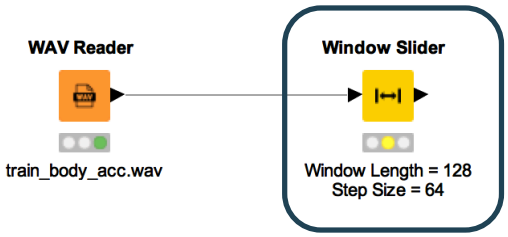
The frequency domain of a square wave looks like a ramp.

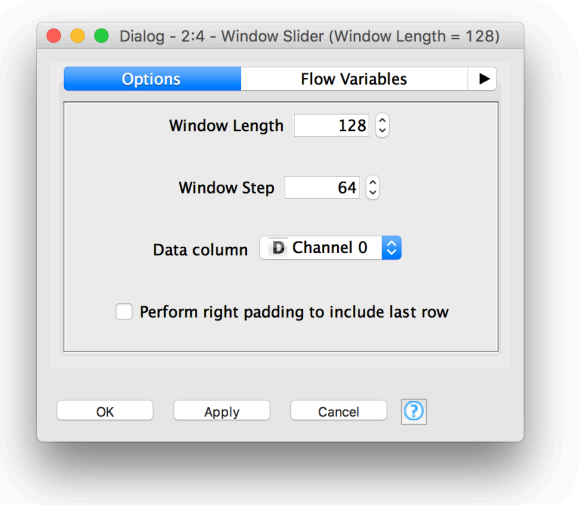


2. Window Slider

- The FFT of a discrete-time signal, which is the Fourier transform of a "windowed" version of the signal, is interpreted as a **sliding-window spectrum**.
- To receive a small "window" of the signal, the window **"slides" across the time series**, one time step at a time.
- It is shown that the signal can be reconstructed from the sampled sliding-window spectrum, i.e., from the values at the points of a certain time-frequency lattice.

2. Window Slider Node

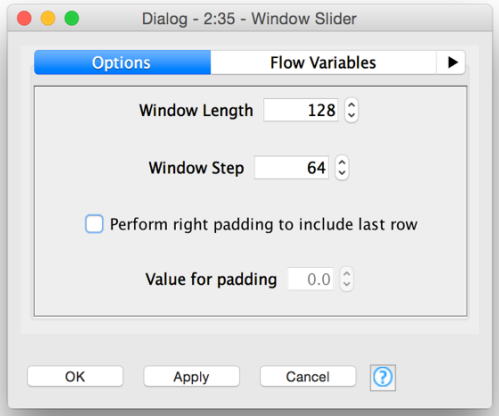




With a sample rate of 50 Hz we will sample fixed-width sliding windows of 2.56 seconds and 50% overlap (128 readings/window).

AdvancedAnalytics .Academy

2. Window Slider Node

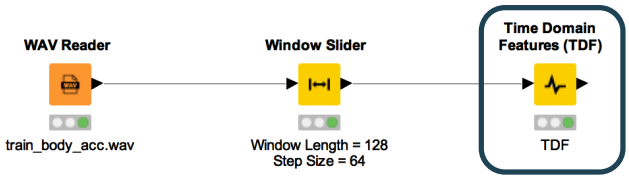


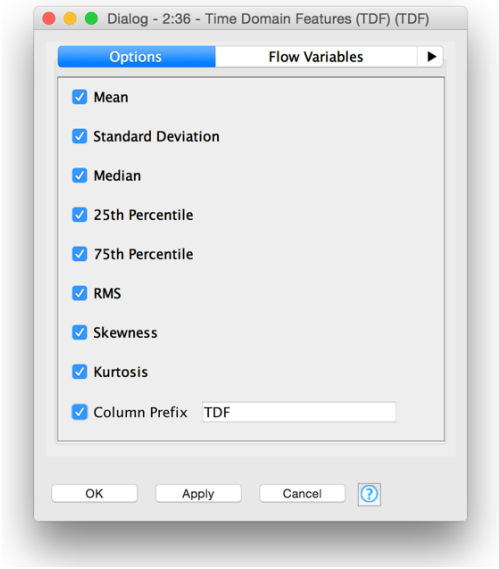
This node creates sliding windows from an input signal which is stored in a table with one column. You can adjust the window length and the step size.

Window Length: Size of a window.
Window Step: Number of timesteps the window moves in one iteration - usually setting this at half the window length is a good idea.
Perform right padding to include last row: Choose if you want to enable padding which compensates for the last window not fitting in the last bit of your data: If you enable padding your signal will be extended by some padding values such that the last window fits into the signal as well. If you don't enable this option the last window will be discarded.
Value for padding:
Choose the value that will be used for padding.

AdvancedAnalytics .Academy

2. Time Domain Features





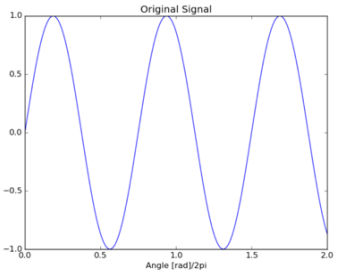
From each window, a vector of features will be obtained by calculating variables from the time domain.

2. Prepare Fast Fourier Transform (Window Function)

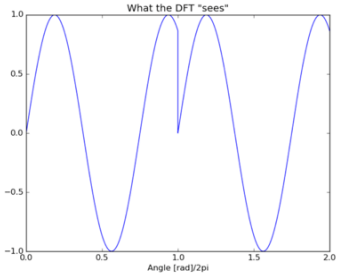
- The FFT computation presumes that the input data repeats over and over.
- This is important when the initial and final values of your data are not the same: the discontinuity causes “leakage” aberrations in the spectrum computed by the FFT.
- "Windowing" smoothes the ends of the data to eliminate these aberrations.
- Windowing premultiplies input data supplied to the FFT with a value that smoothly decreases to zero at each end of data.

2. Prepare Fast Fourier Transform (Window Function)

Original signal



What the FFT sees (Output of Window Slider)

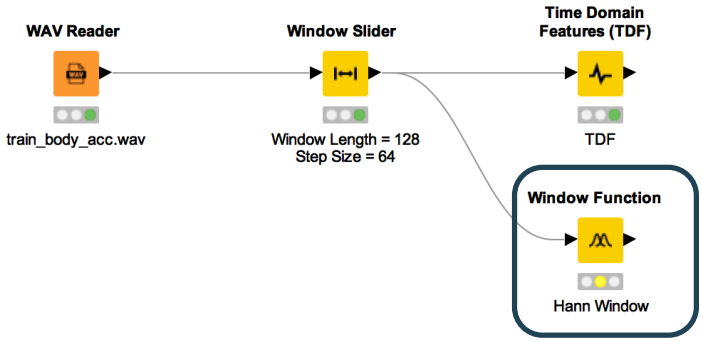


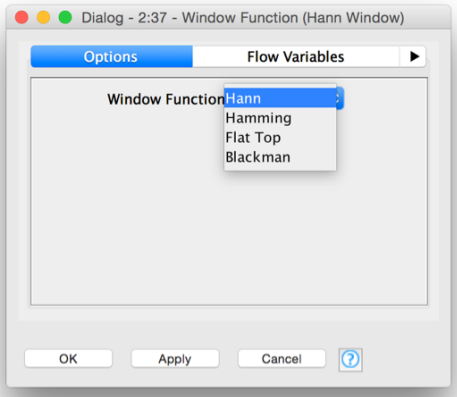
What the FFT sees (Output of Window Function)

2. Window Functions – Rules of thumb

Signal Content	Window
Sine wave or combination of sine waves	Hann
Sine wave (amplitude accuracy is important)	Flat Top
Narrowband random signal (vibration data)	Hann
Broadband random (white noise)	Uniform
Closely spaced sine waves	Uniform, Hamming
Excitation signals (hammer blow)	Force
Response signals	Exponential
Unknown content	Hann
Sine wave or combination of sine waves	Hann
Sine wave (amplitude accuracy is important)	Flat Top
Narrowband random signal (vibration data)	Hann
Two tones with frequencies close but amplitudes very different	Kaiser-Bessel
Two tones with frequencies close and almost equal amplitudes	Uniform
Accurate single tone amplitude measurements	Flat Top

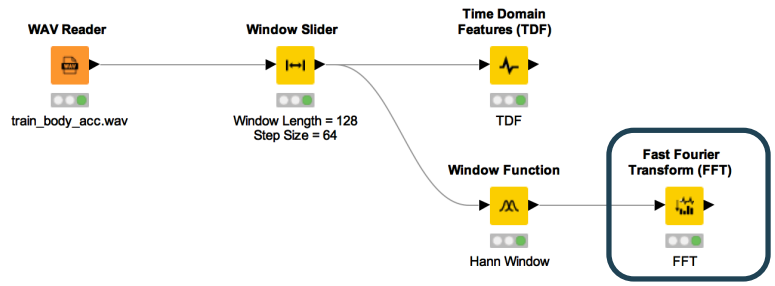
2. Prepare Fast Fourier Transform (Window Function)



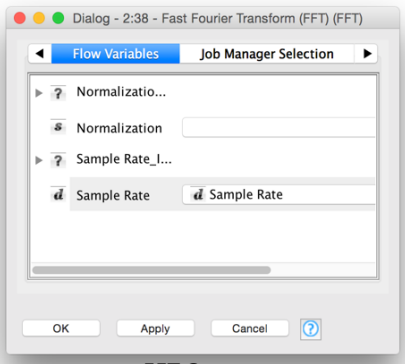


Best Practice: In general, the Hanning (Hann) window is satisfactory in 95 percent of cases. It has good frequency resolution and reduced spectral leakage. If you do not know the nature of the signal but you want to apply a smoothing window, start with the Hann window.

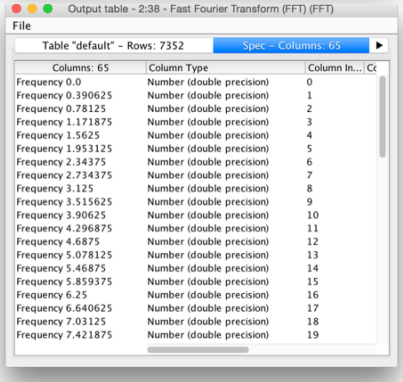
2. Fast Fourier Transform (FFT)



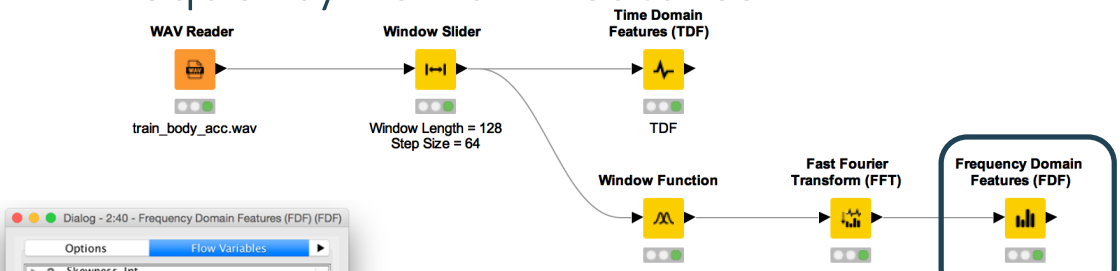
Sample Rate Configuration

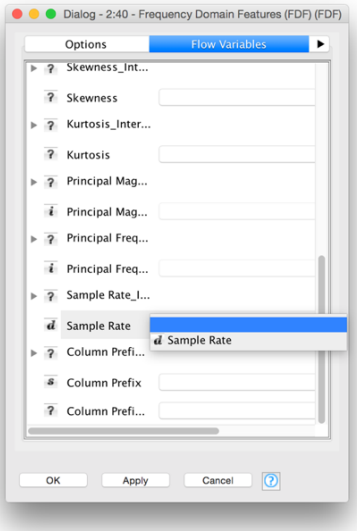


FFT Output



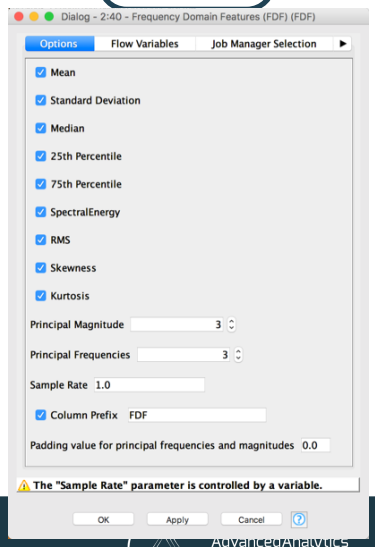
2. Frequency Domain Features



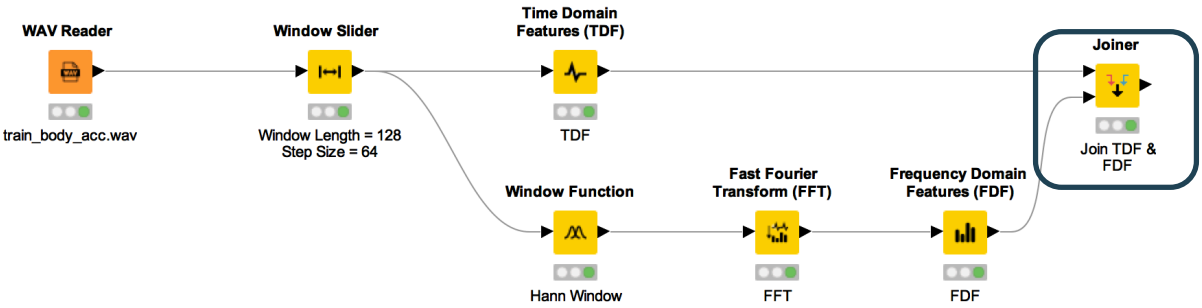


Sample Rate Configuration

From each window, a vector of features will be obtained by calculating variables from the frequency domain.



2. Join Time & Frequency Domain Features



In this example joining of Time Domain and Frequency Domain Features results in 23 Features.

3.
Feature Engineering

3. Feature Engineering

- There are many features that can be extracted from these signals.
- What makes a good feature is that its value has to be similar for signals produced from the same activity and has to be contrasting/different for different user activities.

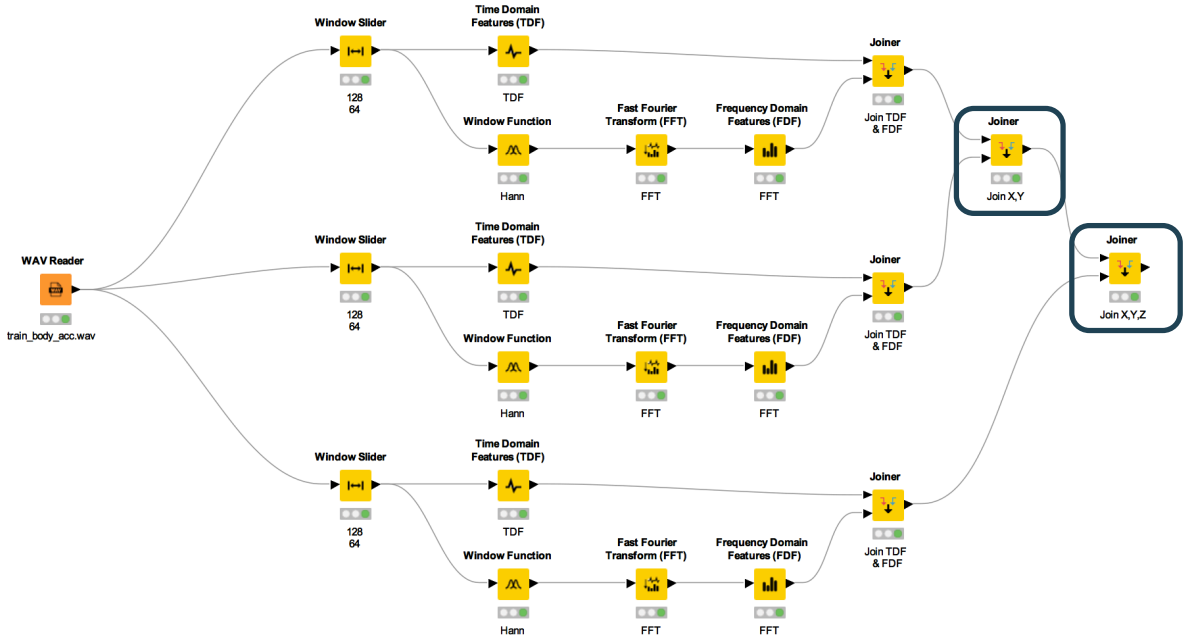
3. Feature Engineering

Number of Features	Description of each feature
9*8 Time Domain Features	Mean, Standard Deviation, Median, 25th Percentile, 75th Percentile, RMS, Skewness, Kurtosis
9*8 Frequency Domain Features	Mean, Standard Deviation, Median, 25th Percentile, 75th Percentile, RMS, Skewness, Kurtosis
9 Frequency Domain Features	Spectral Energy
9*n Frequency Domain Features	n highest peaks of the spectrum (Principal Magnitudes)
9*n Frequency Domain Features	n frequencies with the highest peaks of the spectrum (Principal Frequencies)

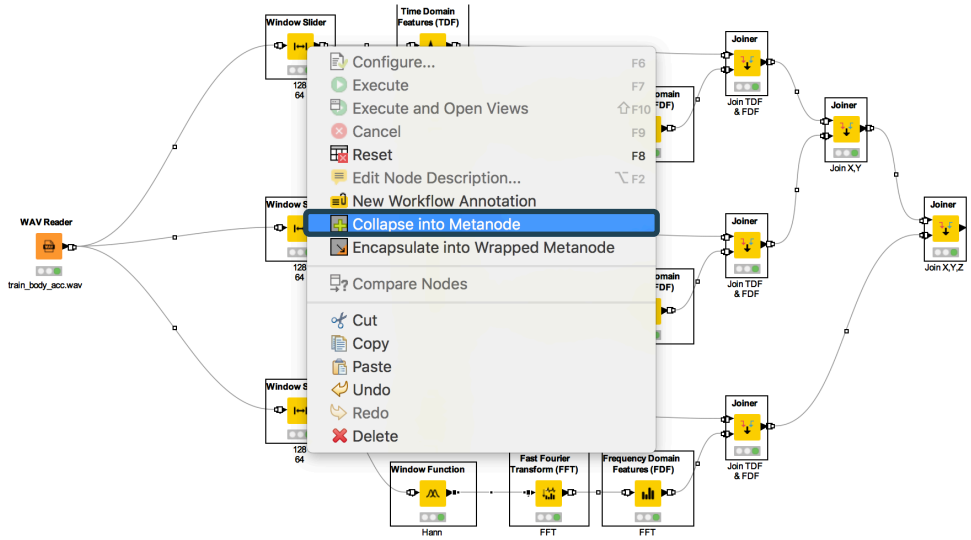
Please note: The features are obtained individually for each signal, since we have 3 datasources with 3 signals coming from 3 axis, we get a minimum of 3 features for every datasource like Mean, or Standard Deviation. Number of features is always a multiple of 9.

3. Prepare features for X-, Y-, Z-Axis

3. Combine & join features for X-, Y-, Z-Axis

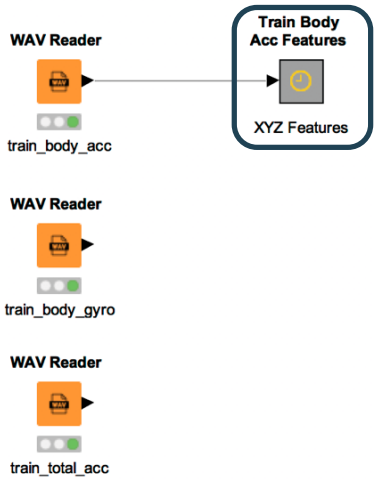


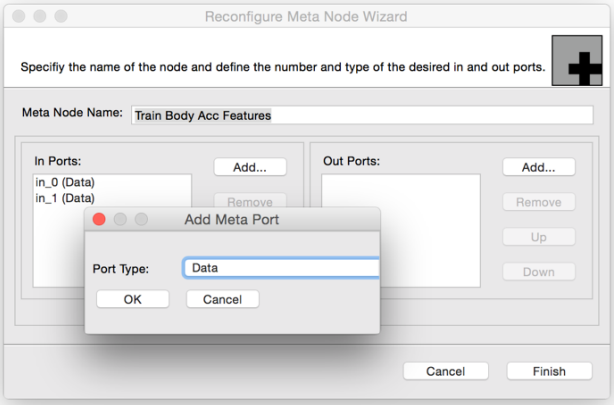
3. Collapse into Meta Node



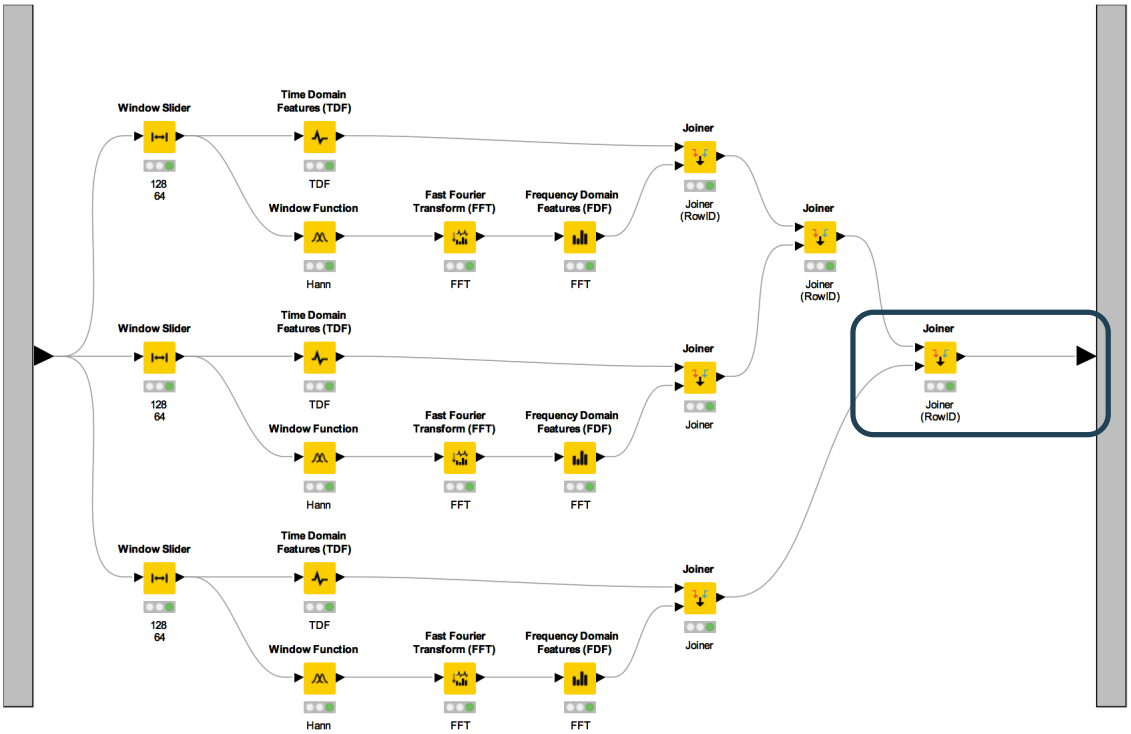
By collapsing the whole feature engineering workflow, the logic of the whole workflow is ready to be reused.

3. Collapse into Meta Node – add data output port



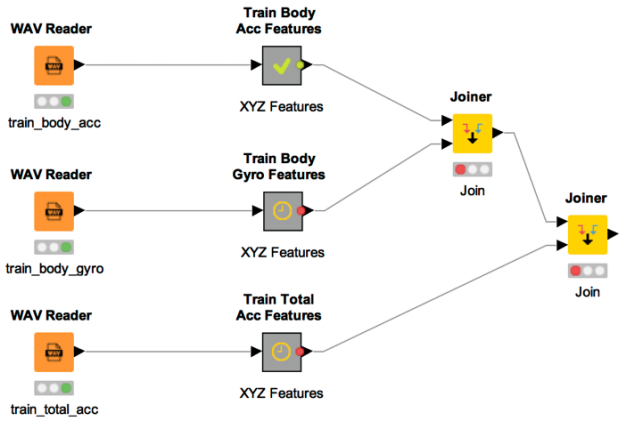


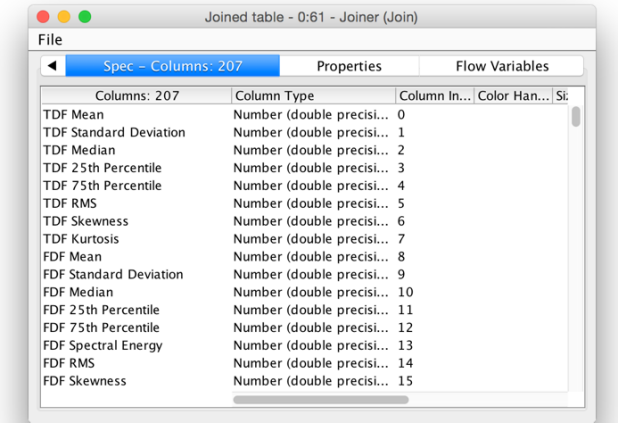
3. Collapse into Meta Node – add data output port



3. Final Feature Engineering workflow

>> The workflow provides 207 Features

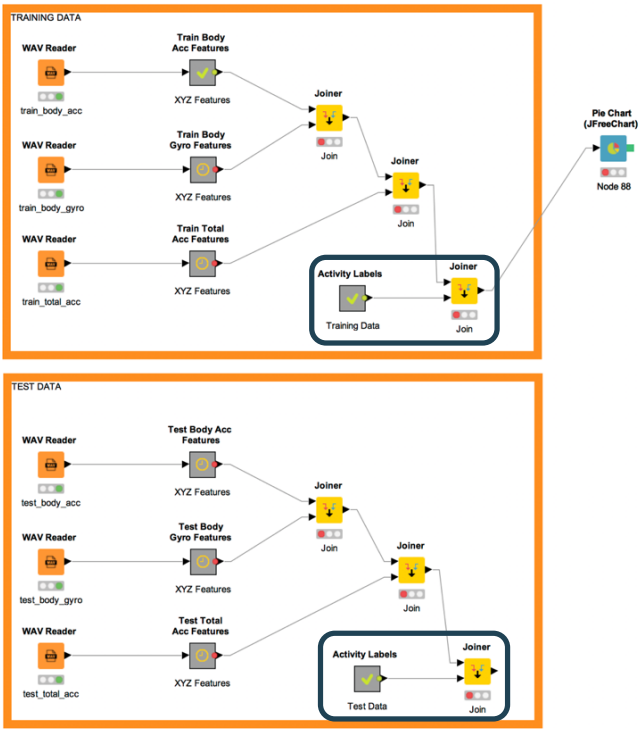


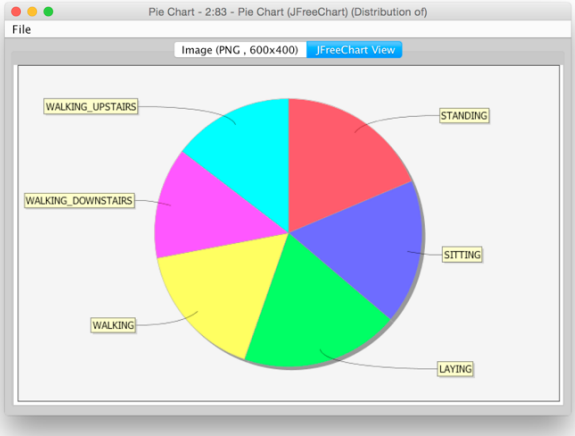


Columns: 207	Column Type	Column In...	Color Han...	Si
TDF Mean	Number (double precisi...	0		
TDF Standard Deviation	Number (double precisi...	1		
TDF Median	Number (double precisi...	2		
TDF 25th Percentile	Number (double precisi...	3		
TDF 75th Percentile	Number (double precisi...	4		
TDF RMS	Number (double precisi...	5		
TDF Skewness	Number (double precisi...	6		
TDF Kurtosis	Number (double precisi...	7		
FDF Mean	Number (double precisi...	8		
FDF Standard Deviation	Number (double precisi...	9		
FDF Median	Number (double precisi...	10		
FDF 25th Percentile	Number (double precisi...	11		
FDF 75th Percentile	Number (double precisi...	12		
FDF Spectral Energy	Number (double precisi...	13		
FDF RMS	Number (double precisi...	14		
FDF Skewness	Number (double precisi...	15		

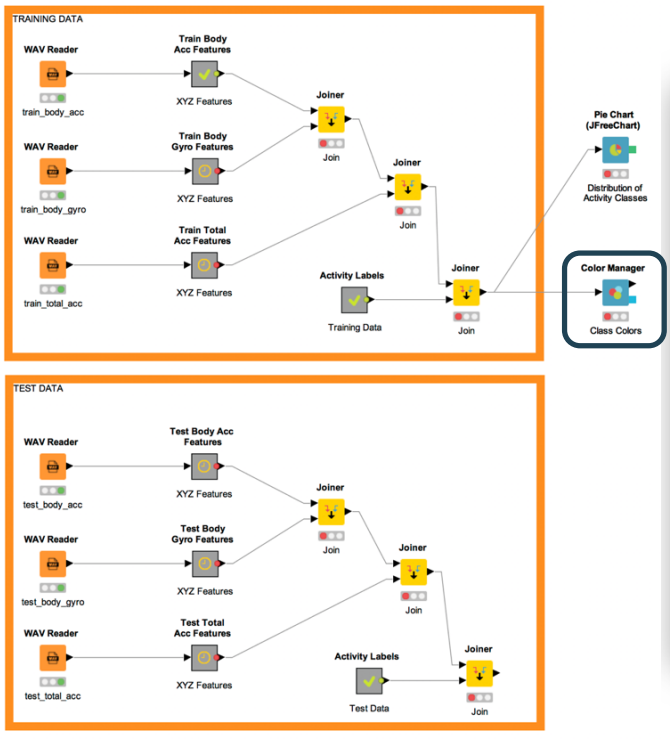
4.
Machine Learning

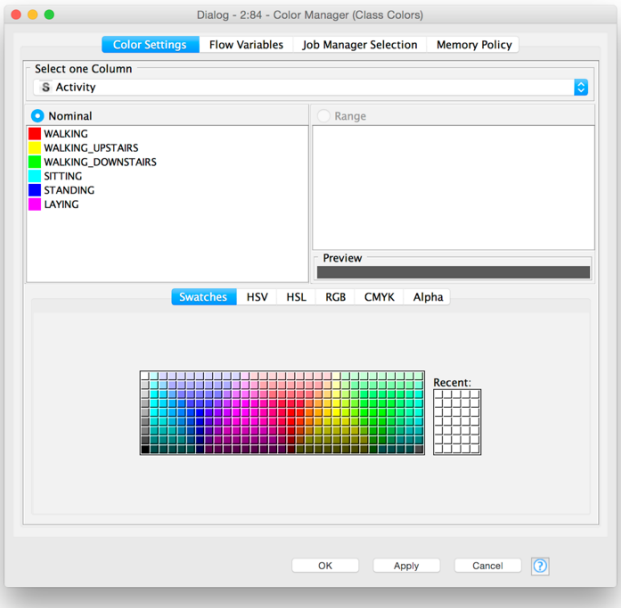
4. Provide Target Classes and Test Data



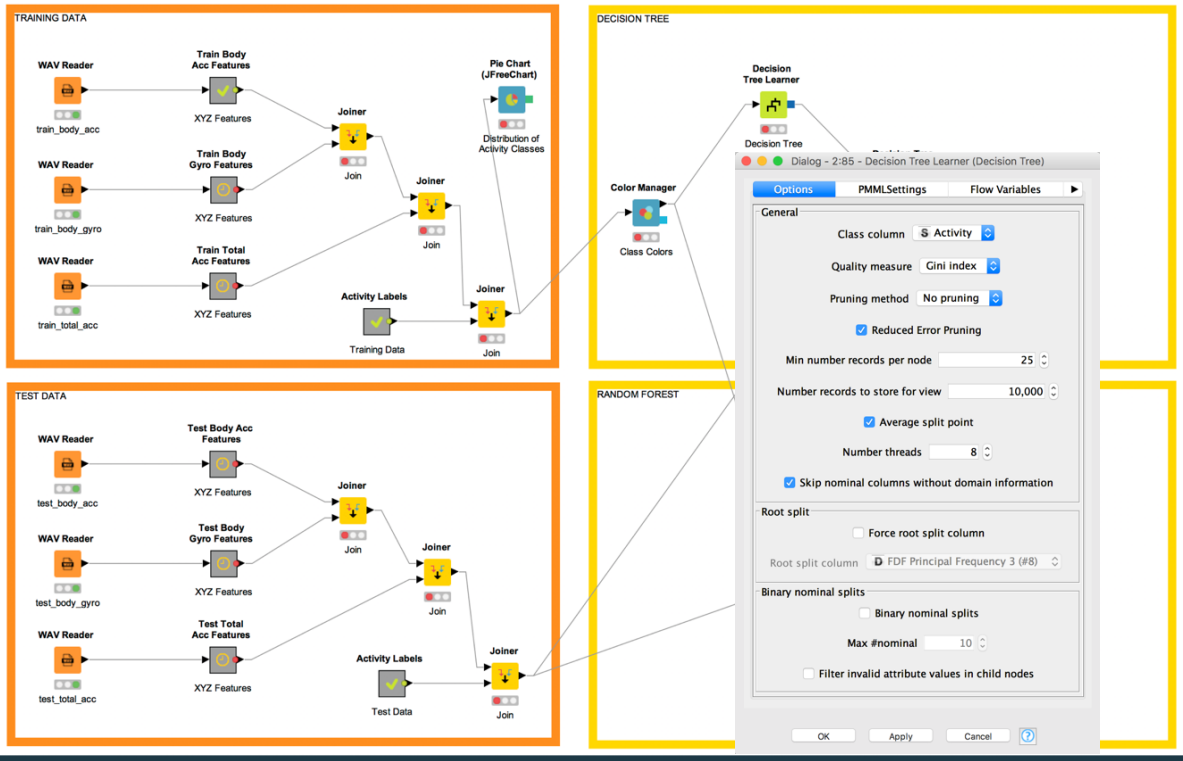


4. Append Color Manager

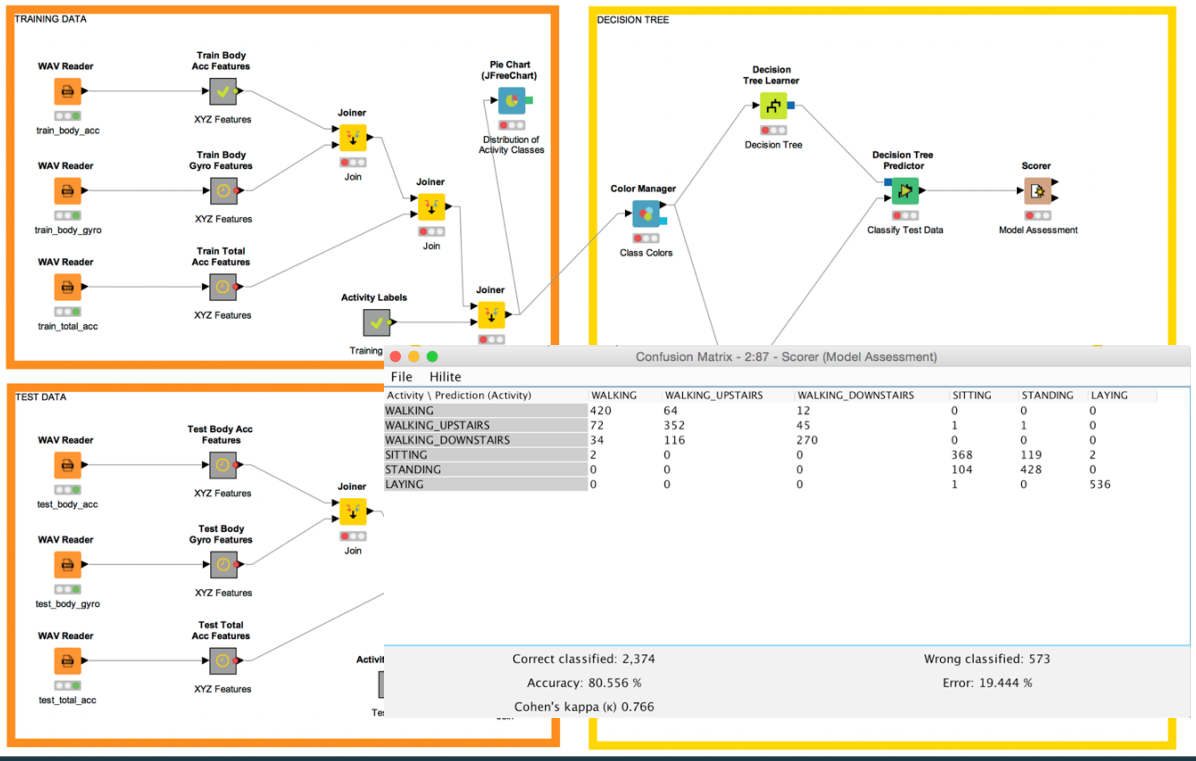




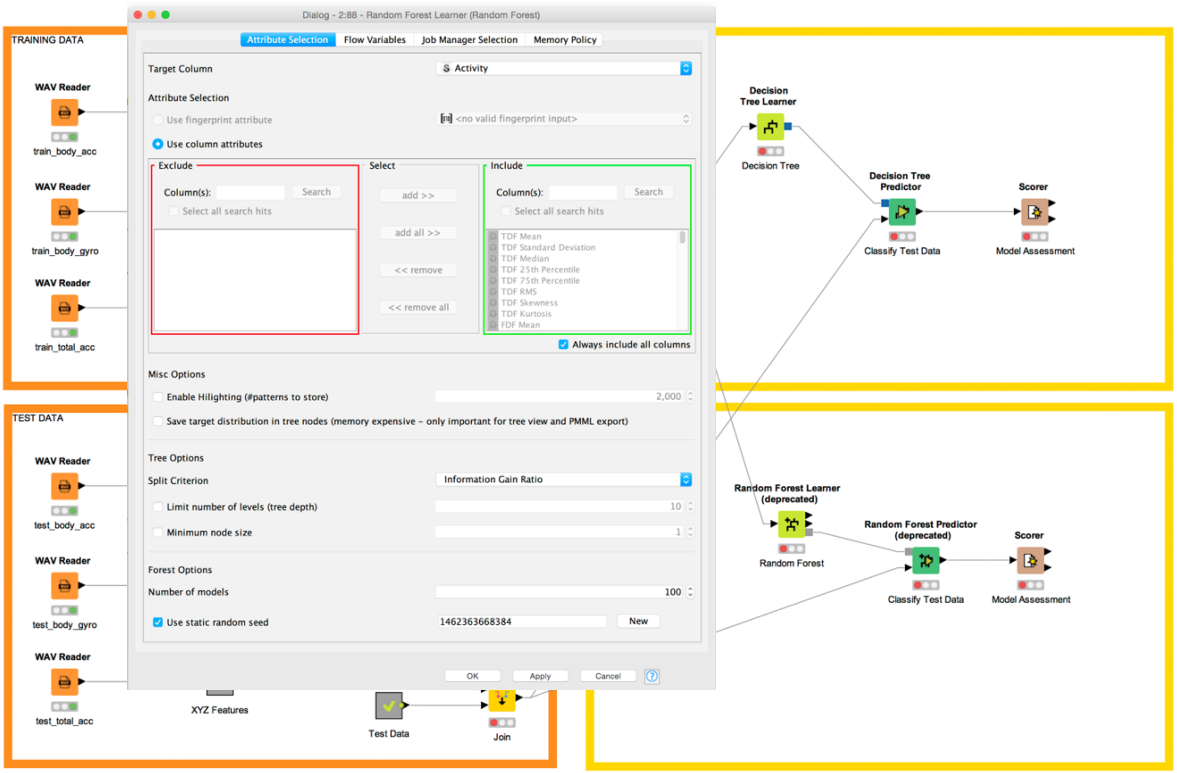
4. Train a Decision Tree Classifier



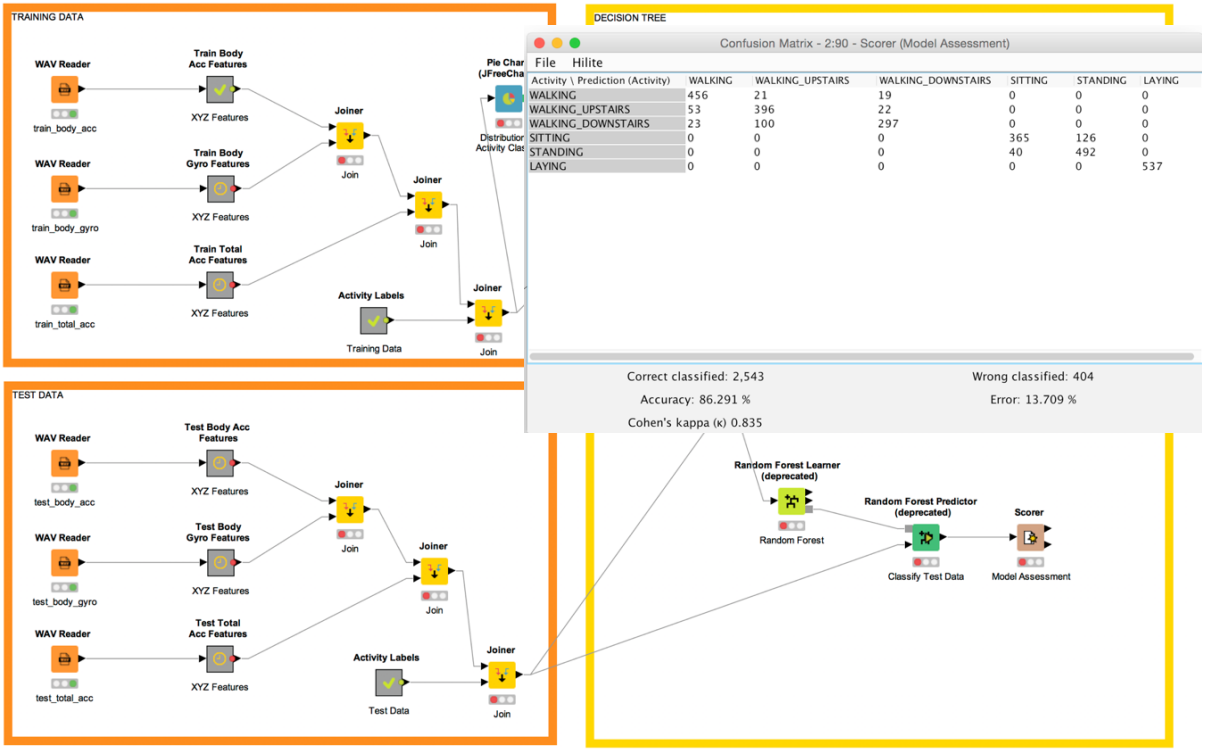
4. Check Classification Results on Test Data

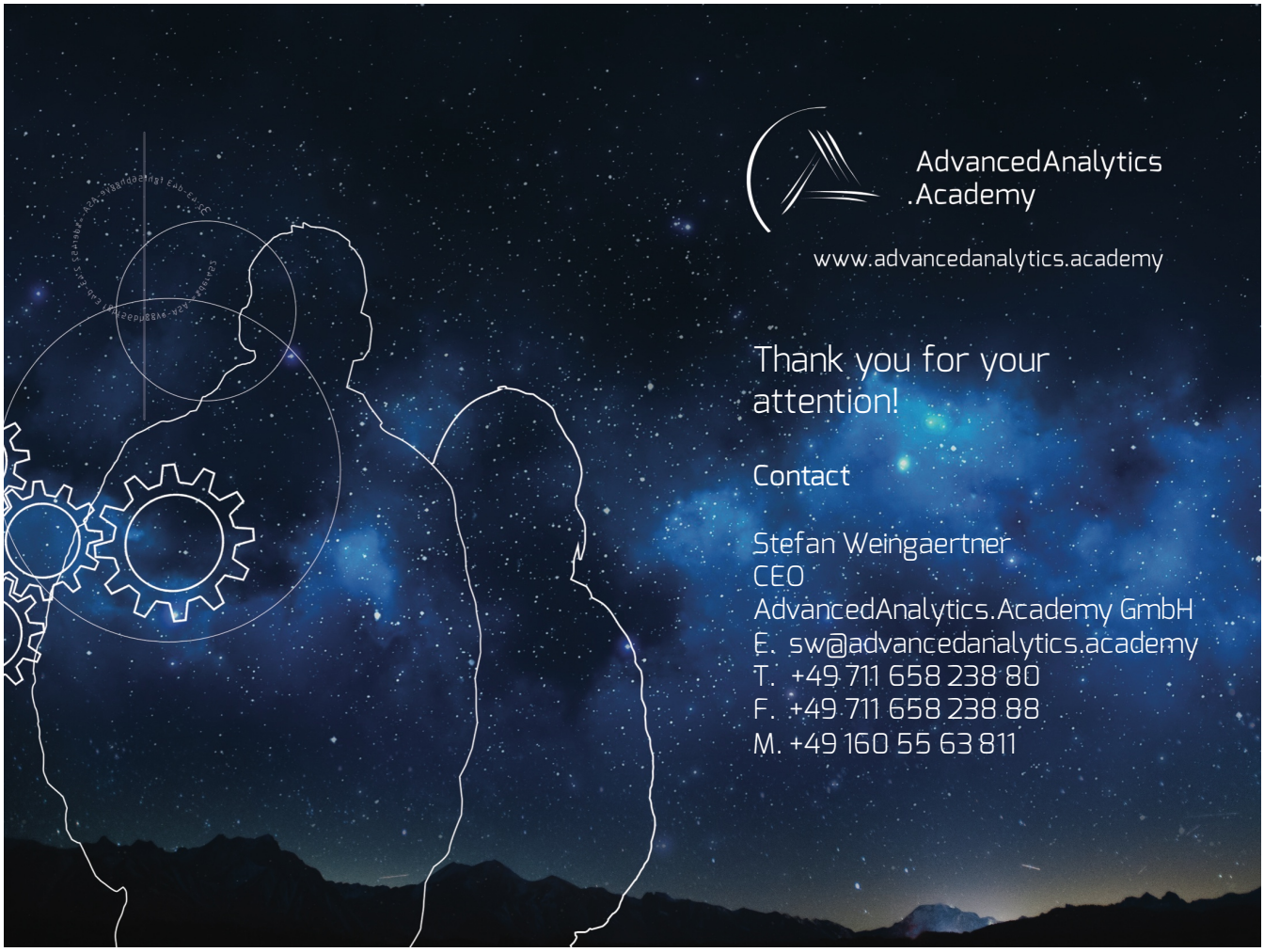


4. Train a Random Forest Classifier



4. Check Classification Results on Test Data




AdvancedAnalytics
.Academy
www.advancedanalytics.academy
Thank you for your attention!
Contact
Stefan Weingaertner
CEO
AdvancedAnalytics.Academy GmbH
E. sw@advancedanalytics.academy
T. +49 711 658 238 80
F. +49 711 658 238 88
M. +49 160 55 63 811