SILICON LABS

works with
BY SILICON LABS

How to Build Edge AI Applications that Work

Examination of common challenges & issues encountered in real-world projects

Chris Knorowski | August 2023

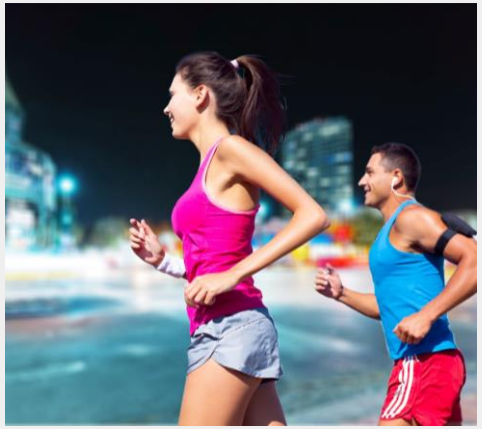
works with
BY SILICON LABS

Our Technology

SensiML™

Our Technology

ACOUSTIC EVENT DETECTION



ACTIVITY RECOGNITION

ANOMALY DETECTION

GESTURE RECOGNITION

KEYWORD SPOTTING

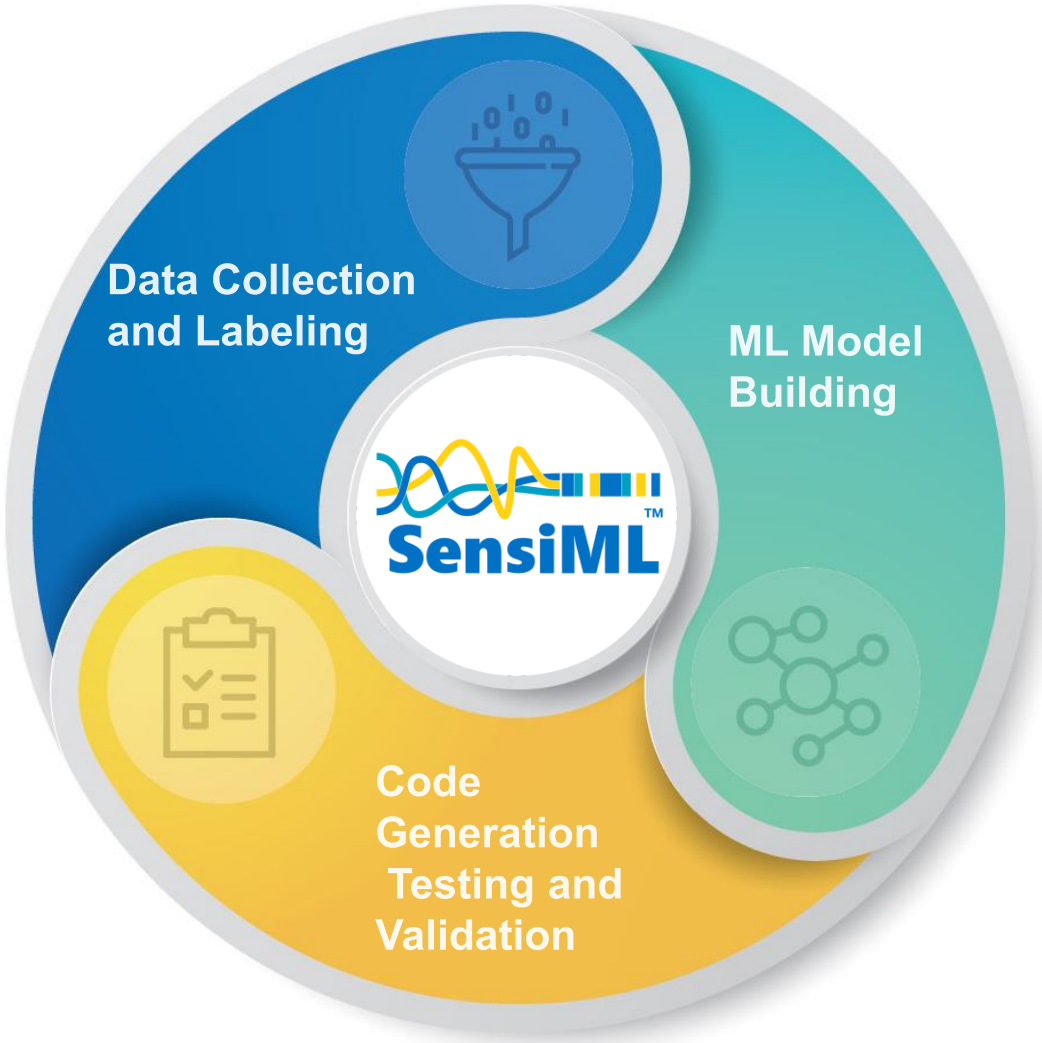
VIBRATION CLASSIFICATION

SensiML Supported Development Kit		
 Silicon Labs xG24 Dev Kit (xG24-DK2601B)	**Processor**	Wireless Gecko EFR32™ Arm® Cortex-M33, 32-bit (EFR32MG24)
	Pre-enabled Sensor Types	TDK ICM-20648 6DoF accel + gyro (data collection firmware), TDK ICS-43434 microphone (left side mic is active in default data collection firmware)
	Additional Available Sensors	Si7021 temp/humidity sensor, VEML6035 light sensor, BMP384 pressure sensor, CCS811 VoC sensor, Si7210 hall-effect sensor
	Available External Sensor Interfaces	UART, I2C, SPI, ADC (12-bit @ 1 Msps,16-bit @ 76.9 ksps)
	Pre-enabled Connectivity	USB, Serial, BLE 5.3 (integrated EFR32 multi-protocol wireless)
	Programming Environment	IDEs: Silicon Labs Simplicity Studio IDE Compilers: Simplicity Studio, gcc, IAR, Keil
	Firmware Flashing	xG24 Dev Kit has built-in programming and debugger via microUSB connection to PC, no separate board or debug cable req'd
	SensiML Knowledge Pack Formats	Binary, Library, C Source
	Useful Links	SensiML Getting Started Guide, Solution Brief, Smart Building IoT Video Workshop, xG24 Smart Door Lock Demo

works with
BY SILICON LABS

Understanding the Machine Learning Workflow

Steps to Build An Edge AI Model



Data Capture Lab

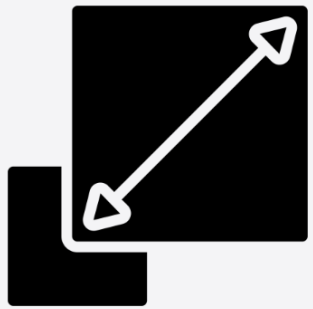
Analytics Studio

Knowledge Pack

works with
BY SILICON LABS

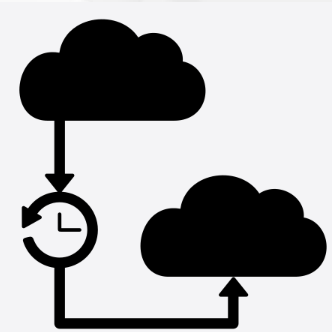
Qualifying and Application

Edge AI: The Right Tool for the Application?



SCALABLE

SECURITY/PRIVACY



LATENCY

REMOTE DEVICES / LIMITED CONNECTIVITY

POWER

ECONOMICAL

works with
BY SILICON LABS

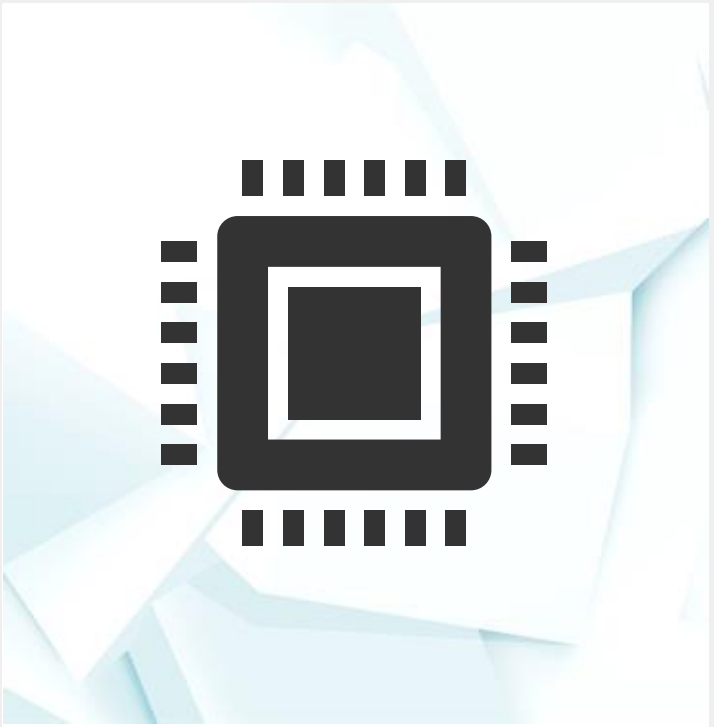
Steps to Build an Application that Uses Edge AI

Edge AI Application Customer Engagement Process

- Customer wants feature X
- Do initial demo of tools/model with internal PoC similar to their use case convince yourselves this can be done and convince them externally that the feature is possible.
- Come up with an SoW to work on a PoC for their specific application
- Come up with data collection protocol/plan for PoC
- Carry out data collection plan with customer device
- Build model
- Integrate model into PoC for test and validation
- Customer happy with PoC, but now wants it in the product

What are the Keys to the Success of an Edge AI Application?

MACHINE LEARNING
DATA SCIENCE
ANALYTICS

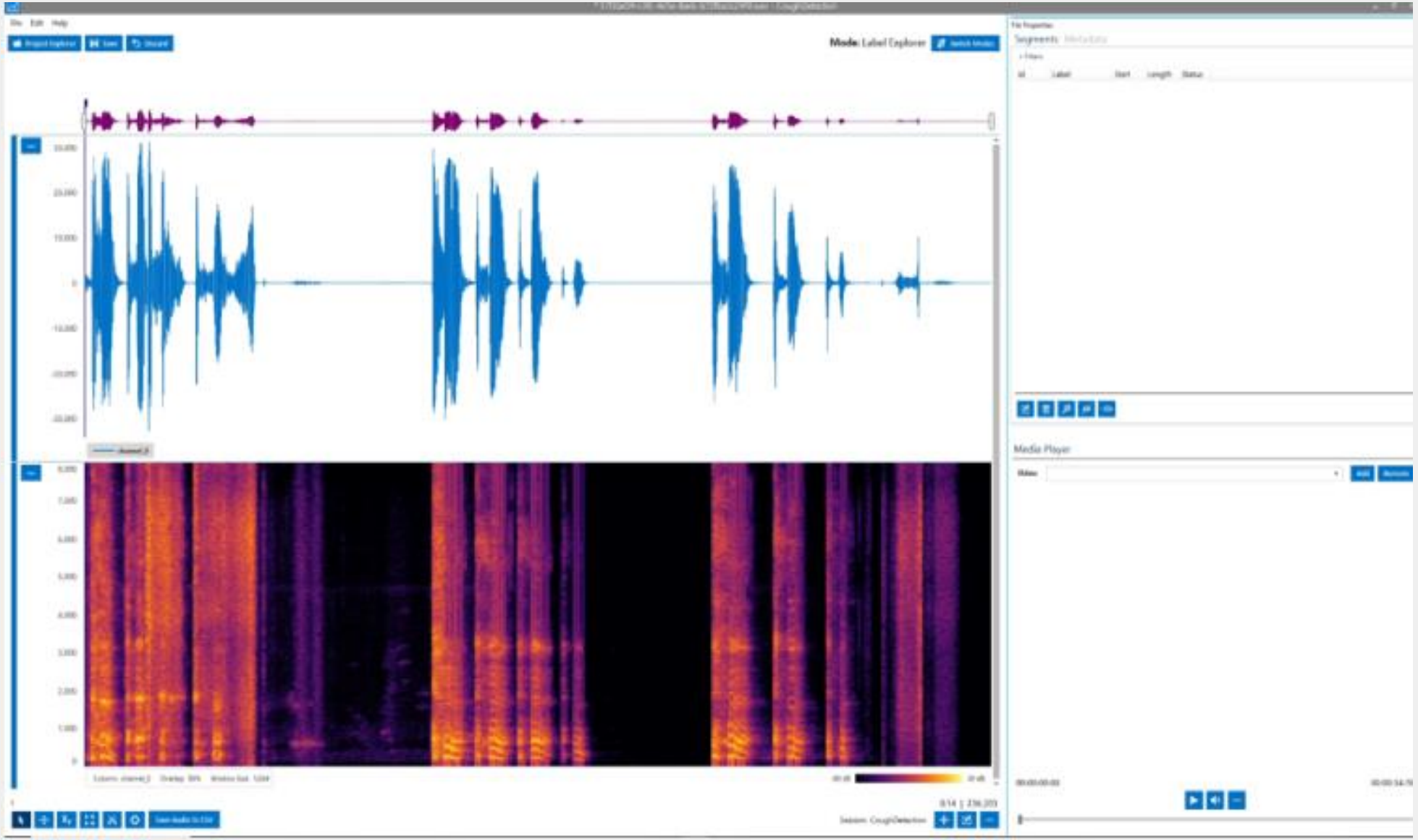


EMBEDDED FIRMWARE
HARDWARE
CONNECTIVITY

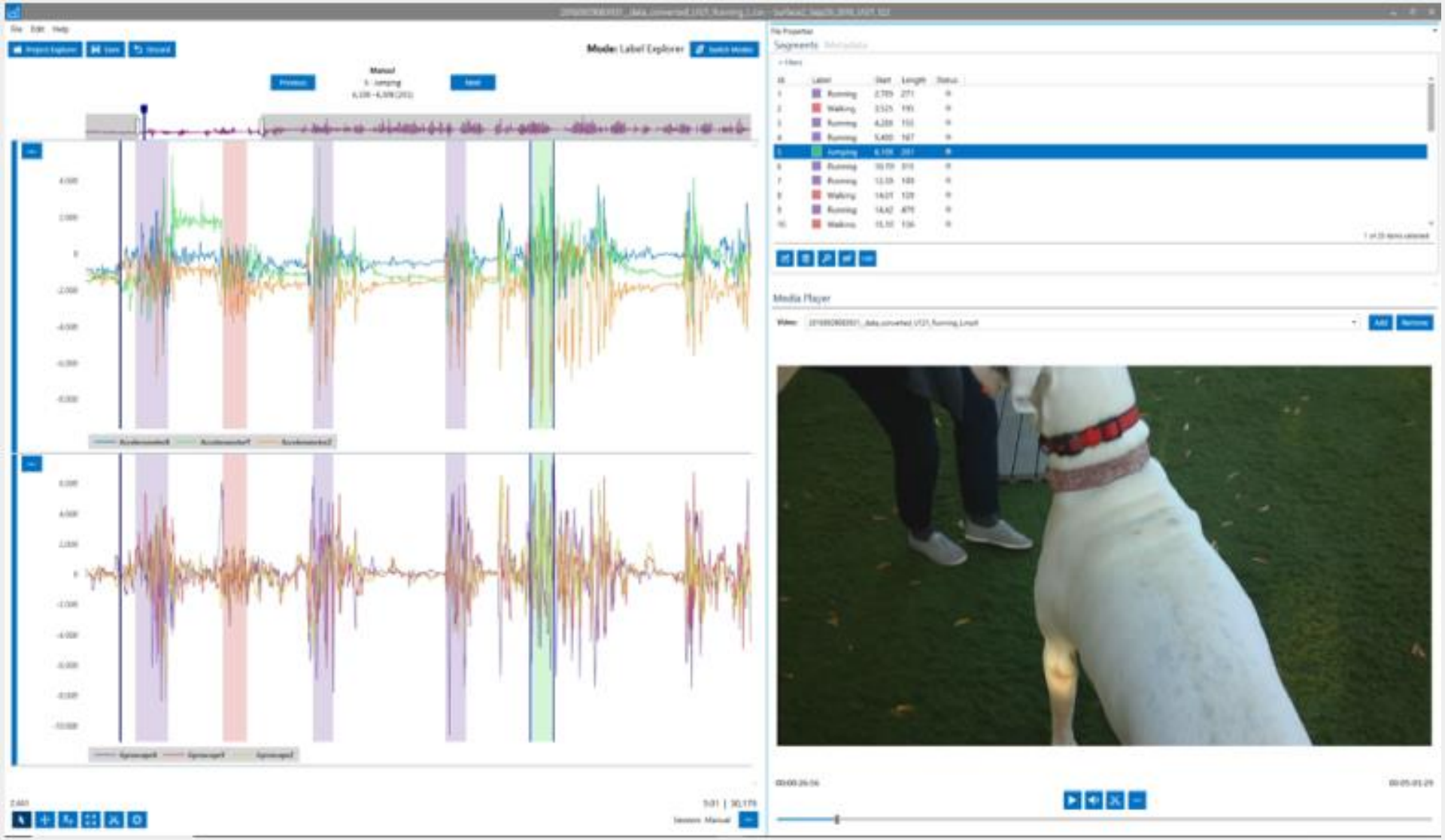
DOMAIN AND
BUSINESS
EXPERTISE

works with
BY SILICON LABS

Data Set Creation


Mode: Label Explorer
Media Player

Understanding the Sensors for your Application

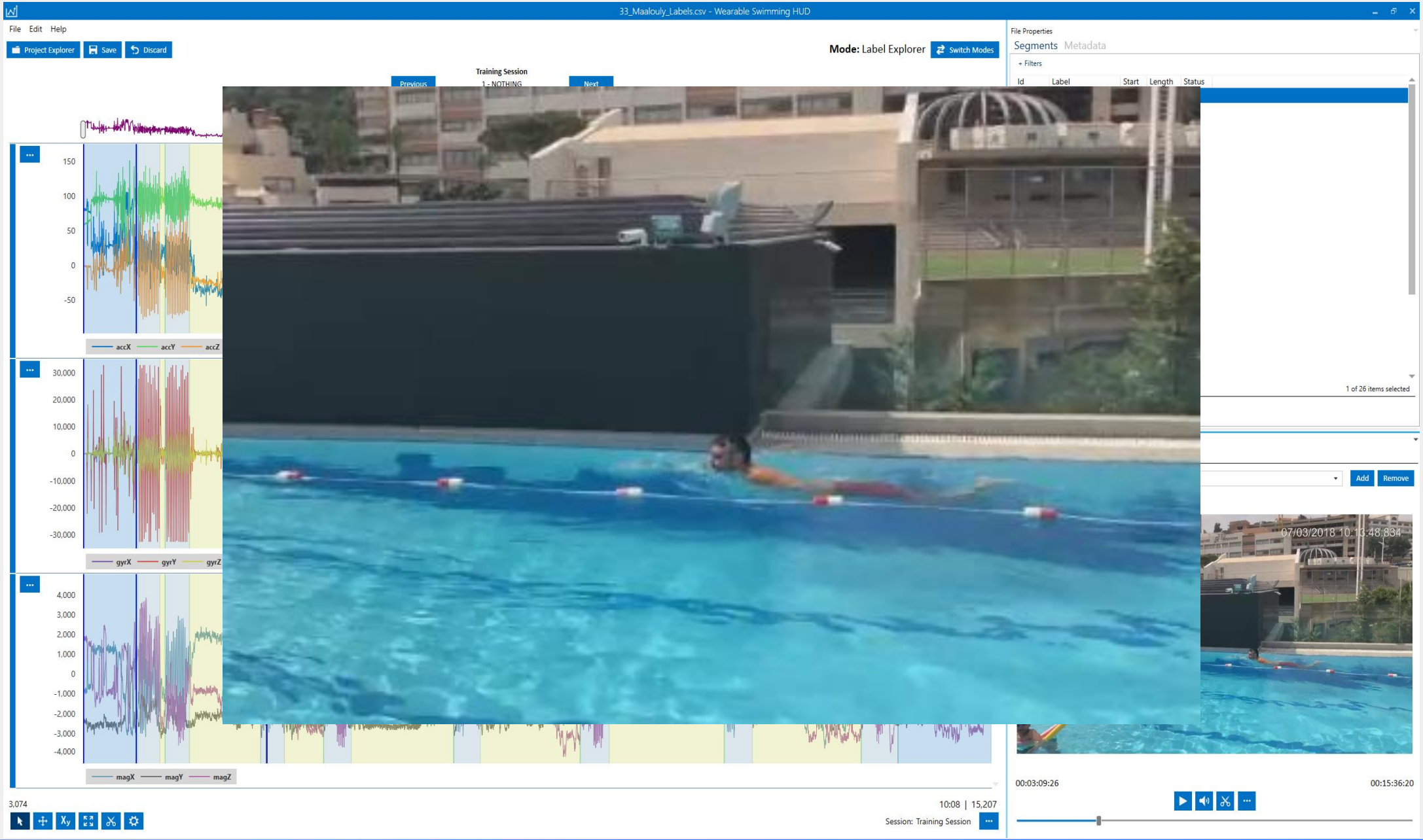

Mode: Label Explorer
Previous
Next
Segments
Media Player

works with
BY SILICON LABS

Defining the Model Scope and Context

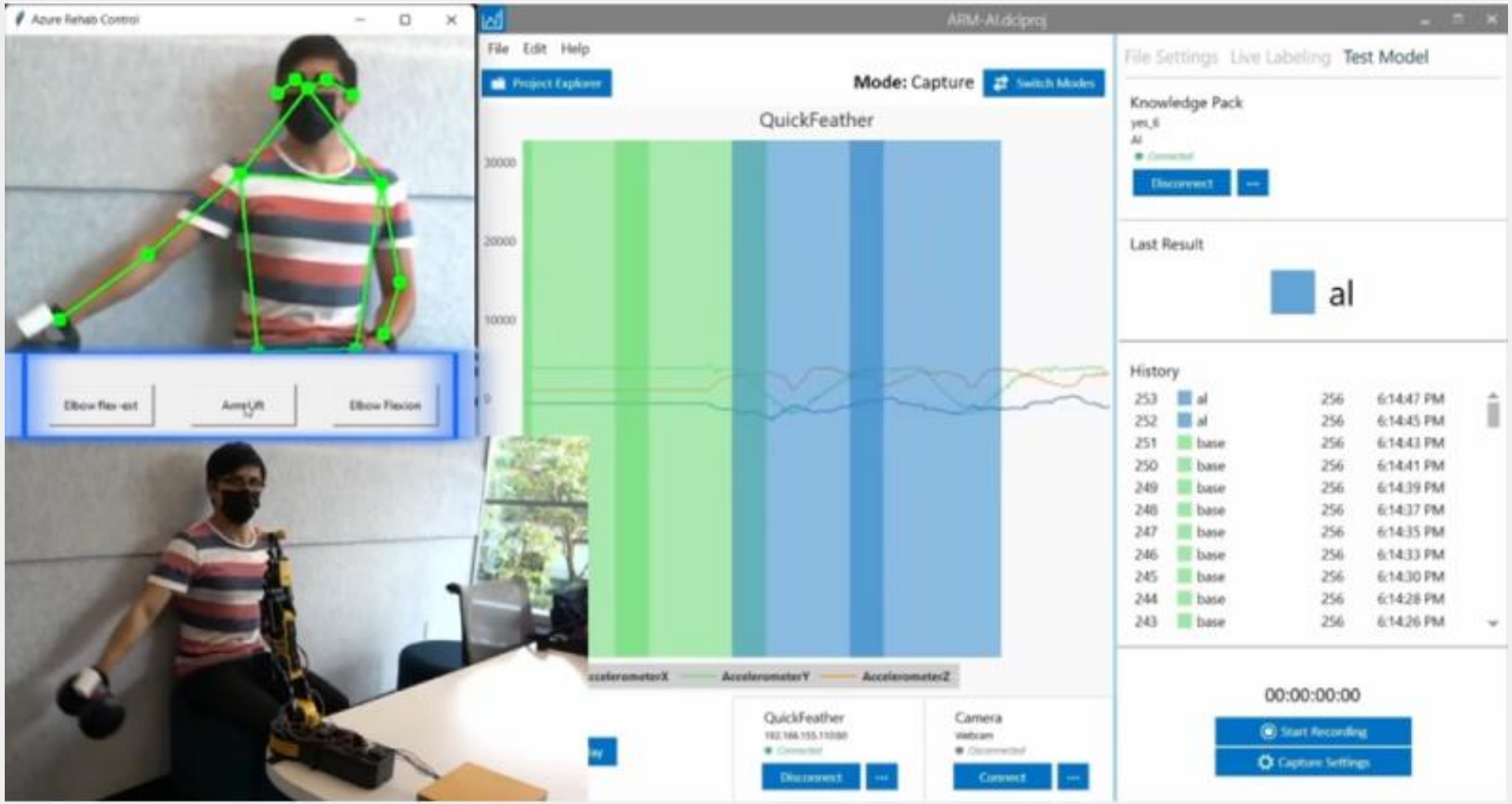

1002_Hind_Labels.csv - Wearable Swimming HUD
File Edit Help
Project Explorer
Save
Discard
Mode: Label Explorer
Switch Modes
Training Session
20 - BREAST
3,311 - 3,656 (346)
Previous
Next
accX accY accZ
gyrX gyrY gyrZ
magX magY magZ
File Properties
Segments Metadata
+ Filters
Id Label Start Length Status
2 START 1,107 2
3 UNDERWATER 1,109 93
4 FLY 1,203 241
5 OPEN 1,444 57
6 UNDERWATER 1,502 84
7 FLY 1,587 245
8 OPEN 1,832 50
9 UNDERWATER 1,883 35
10 BACK 1,919 339
11 FLIP 2,258 34
12 UNDERWATER 2,293 30
13 BACK 2,324 365
14 REST 2,689 89
15 START 2,779 17
16 UNDERWATER 2,797 86
17 BREAST 2,884 326
18 OPEN 3,210 50
19 UNDERWATER 3,261 49
20 BREAST 3,311 346
1 of 28 items selected
Media Player
Video: 1002_Hind_Labels.mp4
Add
Remove
00:02:31:88
00:04:40:12
2,631
3:14 | 4,854
Session: Training Session


33_Maalouly_Labels.csv - Wearable Swimming HUD
File Edit Help
Project Explorer
Save
Discard
Mode: Label Explorer
Switch Modes
Training Session
1 - NOTHING
0 - 895 (896)
Previous
Next
accX accY accZ
gyrX gyrY gyrZ
magX magY magZ
3,074
10:08 | 15,207
Session: Training Session
File Properties
Segments Metadata
+ Filters
Id Label Start Length Status
1 NOTHING 0 896
2 START 896 13
3 FREE 910 373
4 REST 1,283 78
5 START 1,362 7
6 FREE 1,370 406
7 REST 1,776 1,191
8 START 2,967 4
9 FREE 2,972 402
10 REST 3,374 648
11 START 4,022 9
12 FREE 4,032 458
13 REST 4,490 1,707
14 START 6,197 7
15 FREE 6,205 448
16 REST 6,653 1,665
17 START 8,318 8
18 FREE 8,327 487
19 REST 8,814 1,919
1 of 26 items selected
Media Player
Video: 33_Maalouly_Labels.mp4
Add
Remove
07/03/2018 10:13:48.834
00:03:09:26
00:15:36:20


33_Maalouly_Labels.csv - Wearable Swimming HUD
File Edit Help
Project Explorer
Save
Discard
Mode: Label Explorer
Switch Modes
Training Session
1 - NOTHING
Previous
Next
accX accY accZ
gyrX gyrY gyrZ
magX magY magZ
File Properties
Segments
Metadata
+ Filters
Id
Label
Start
Length
Status
1 of 26 items selected
Add
Remove
07/03/2018 10:13:48.834
00:03:09:26
00:15:36:20
3,074
10:08 | 15,207
Session: Training Session

works with
BY SILICON LABS

Rapid Model Test and Validation Framework


Azure Rehab Control
Elbow flex-ext
Elbow Flexion
ARM-AI.ddproj
File Edit Help
Project Explorer
Mode: Capture
Switch Modes
QuickFeather
30000
20000
10000
0
AccelerometerY
AccelerometerZ
QuickFeather
Disconnect
Camera
Webcam
Connect
File Settings
Live Labeling
Test Model
Knowledge Pack
yes_ti
AI
Disconnect
Last Result
al
History
253 al 256 6:14:47 PM
252 al 256 6:14:45 PM
251 base 256 6:14:43 PM
250 base 256 6:14:41 PM
249 base 256 6:14:39 PM
248 base 256 6:14:37 PM
247 base 256 6:14:35 PM
246 base 256 6:14:33 PM
245 base 256 6:14:30 PM
244 base 256 6:14:28 PM
243 base 256 6:14:26 PM
00:00:00:00
Start Recording
Capture Settings

works with
BY SILICON LABS

Training Neural Networks For Edge Devices

Training Neural Networks for EdgeAI

1. **Collect your data** set, then use **augmentation** to tailor it to the specific environment the model will operate in as well as expand the number of examples of each event

2. Select a **foundation model** that was trained on a large corpus of keywords and acoustic sounds and has the appropriate architecture

3. Use **transfer learning** to train your new model to detect your specific keywords along with **feature augmentation** to prevent overfitting

4. Use **quantization aware training** to tune the model for an edge device

5. Use **post training quantization** to quantize the model and make it suitable for deploying at the edge

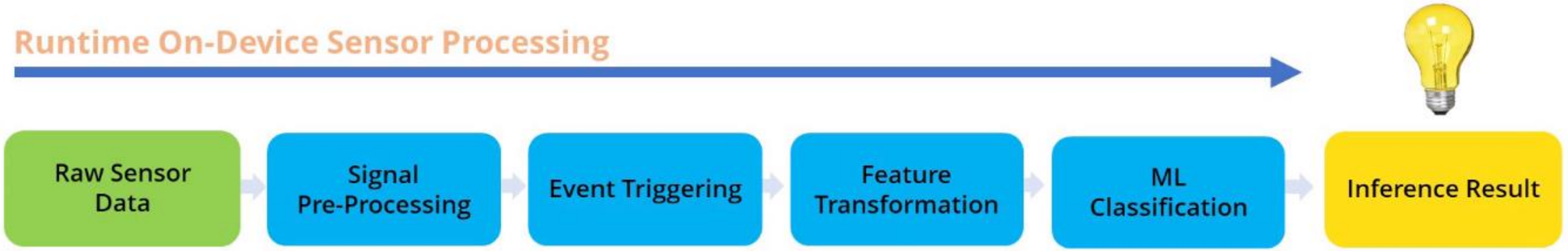
works with
BY SILICON LABS

Building a Good Test Data Set

works with
BY SILICON LABS

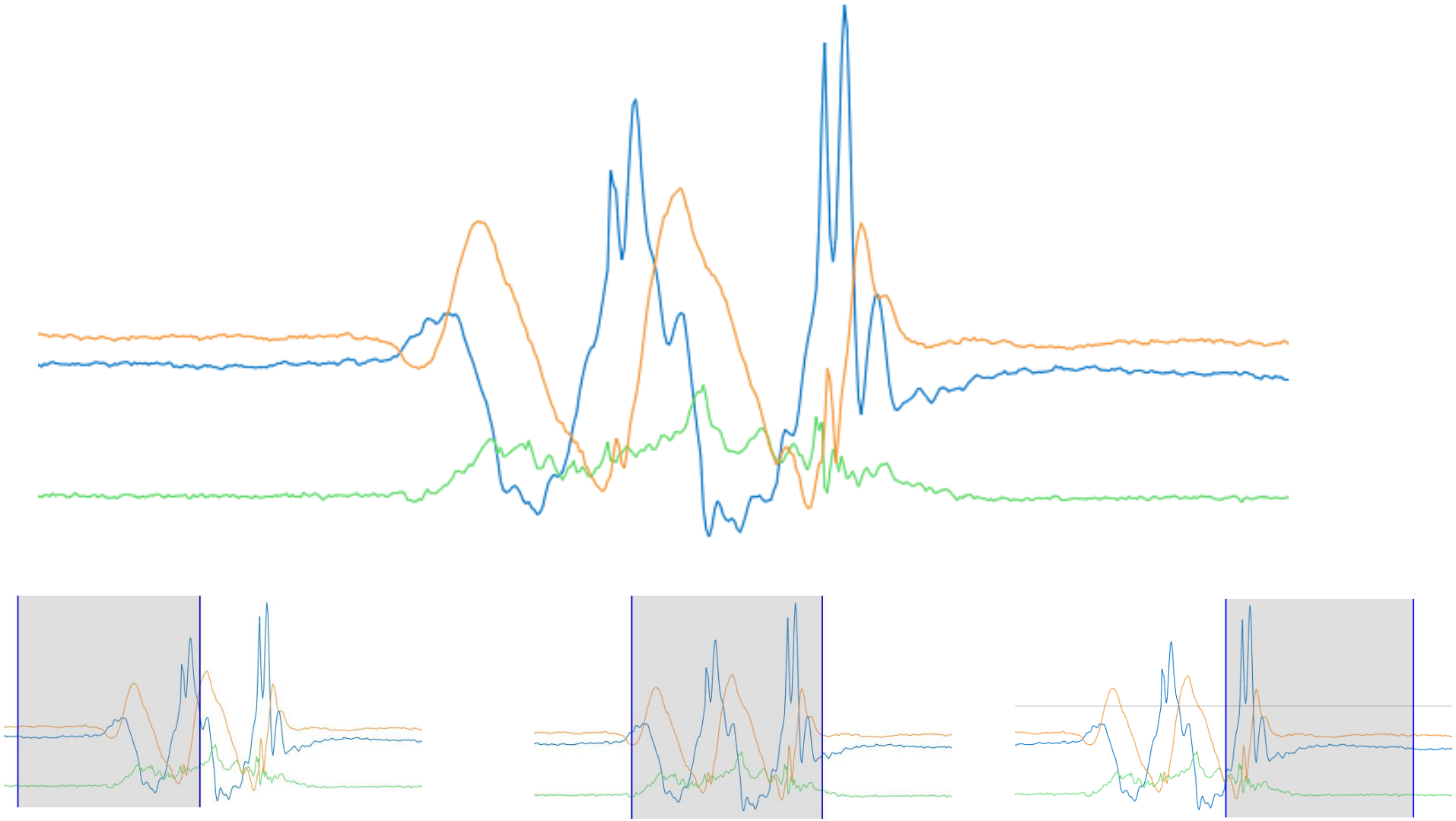
DSP Preprocessing

KnowledgePack – DSP Pipeline and Classifier Firmware



- Segmentation can help reduce false positives and class confusion
- Segmentation prevents running classifier unnecessarily, reducing power consumption
- Feature extraction can make a complicated problem simple
- Features with good class separation reduce the classifier complexity
- Less features reduces the model complexity
- Less model complexity -> lower latency and memory required

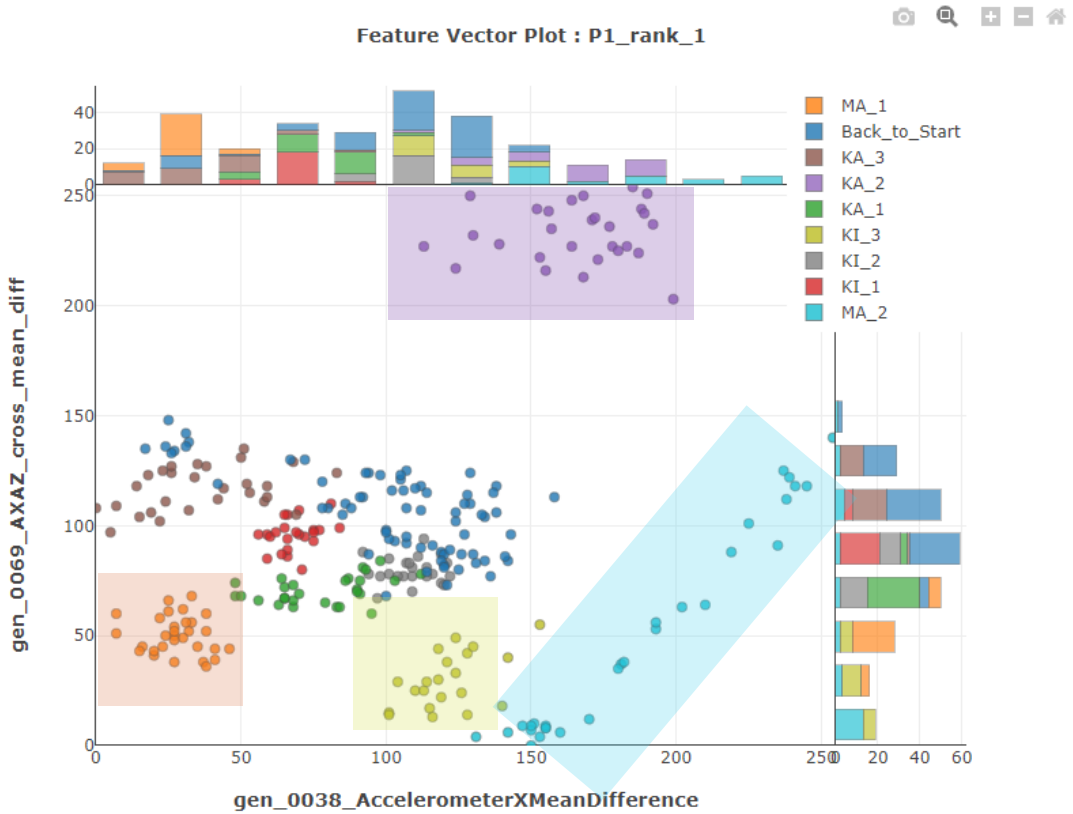
Event Triggering and Segmentation



Feature Transformation/Extraction

Id	Label	Start	Length
2	KA_1	273	145
3	KA_2	547	161
4	KA_3	808	154
5	Back_to_Start	1,063	144
6	KA_1	1,460	209
7	KA_2	1,762	150
8	KA_3	1,995	124

~150x3=450



works with
BY SILICON LABS

Use The Right Classifier For Your Application

AutoML Hyperparameter Search with Cross-Fold Validation

- **Linear Regression models - << 1k**
- **KNN models 1 K < 50 K**
- **Decision Tree Models 10 K < 50 K**
- **Boosted Decision Tree Models 25 K < 200 K**
- **Bonsai Decision Tree Models 10K < 25 K**
- **TensorFlow Models 25 K < 1 MB**

AutoML Rapidly searches across all classifiers

89.08	PME	RBF with Neuron Allocation Optimization	19	3460	91.30	88.87	0
71.09	PME	RBF with Neuron Allocation Optimization	3	3456	73.91	73.26	1
66.82	PME	Hierarchical Clustering with Neuron Optimization	33	3460	78.26	75.66	1
48.33	Decision Tree Ensemble	Random Forest	19	3460	64.35	58.89	1
43.47	Decision Tree Ensemble	Random Forest	4	3456	58.26	51.39	1
	TensorFlow Lite for Microcontrollers	Train Fully Connected Neural Network	12	3460	5.22		1
28.14	PME	Hierarchical Clustering with Neuron Optimization	19	3460	50.43	42.22	1
	TensorFlow Lite for Microcontrollers	Train Fully Connected Neural Network	49	3460	5.22		1
76.09	Decision Tree Ensemble	Random Forest	16	3460	85.22	80.37	1

Questions?

Visit us at https://sensiml.com

Chris Knorowski | chris.Knorowski@sensiml.com | August 2023