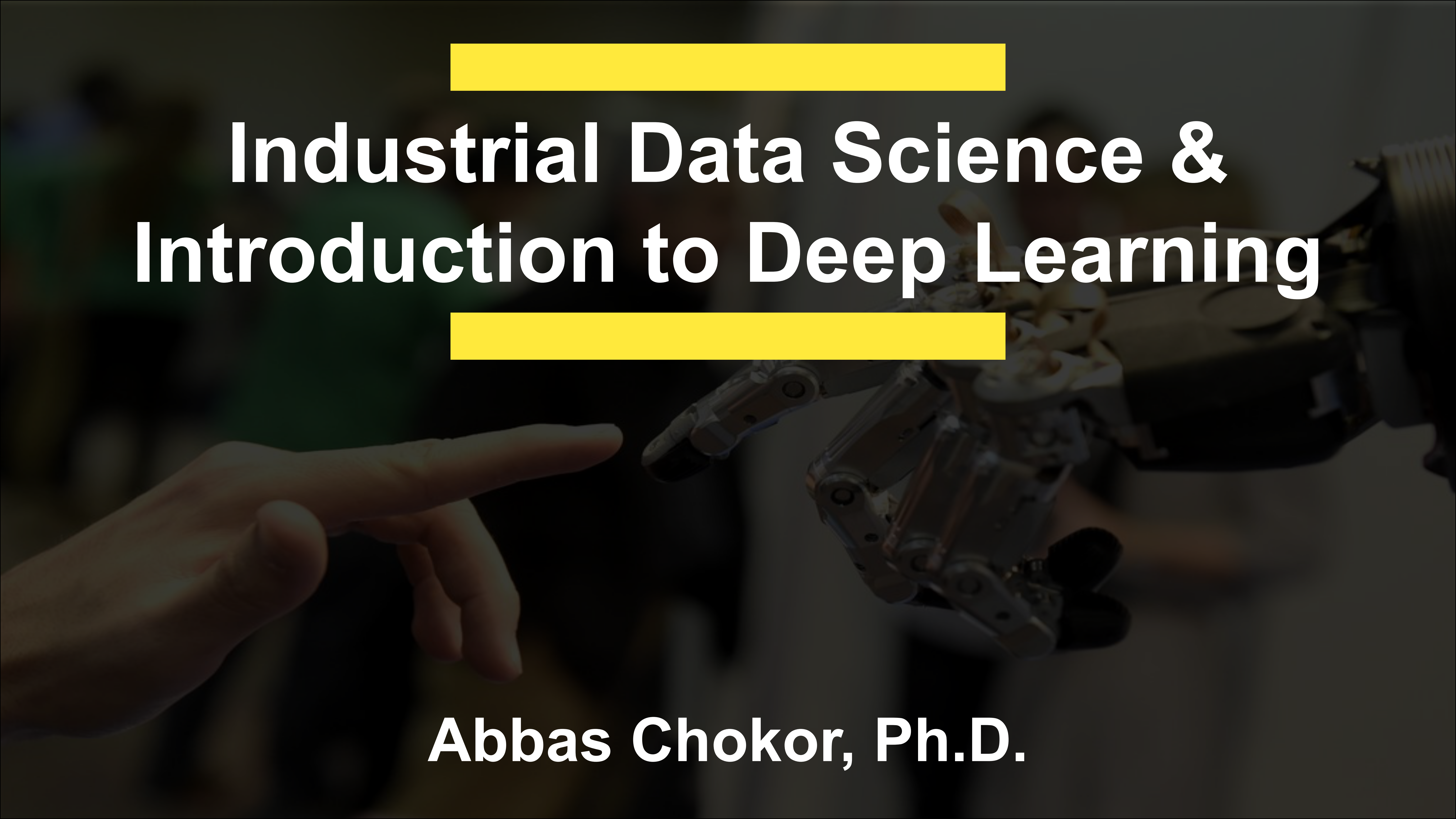
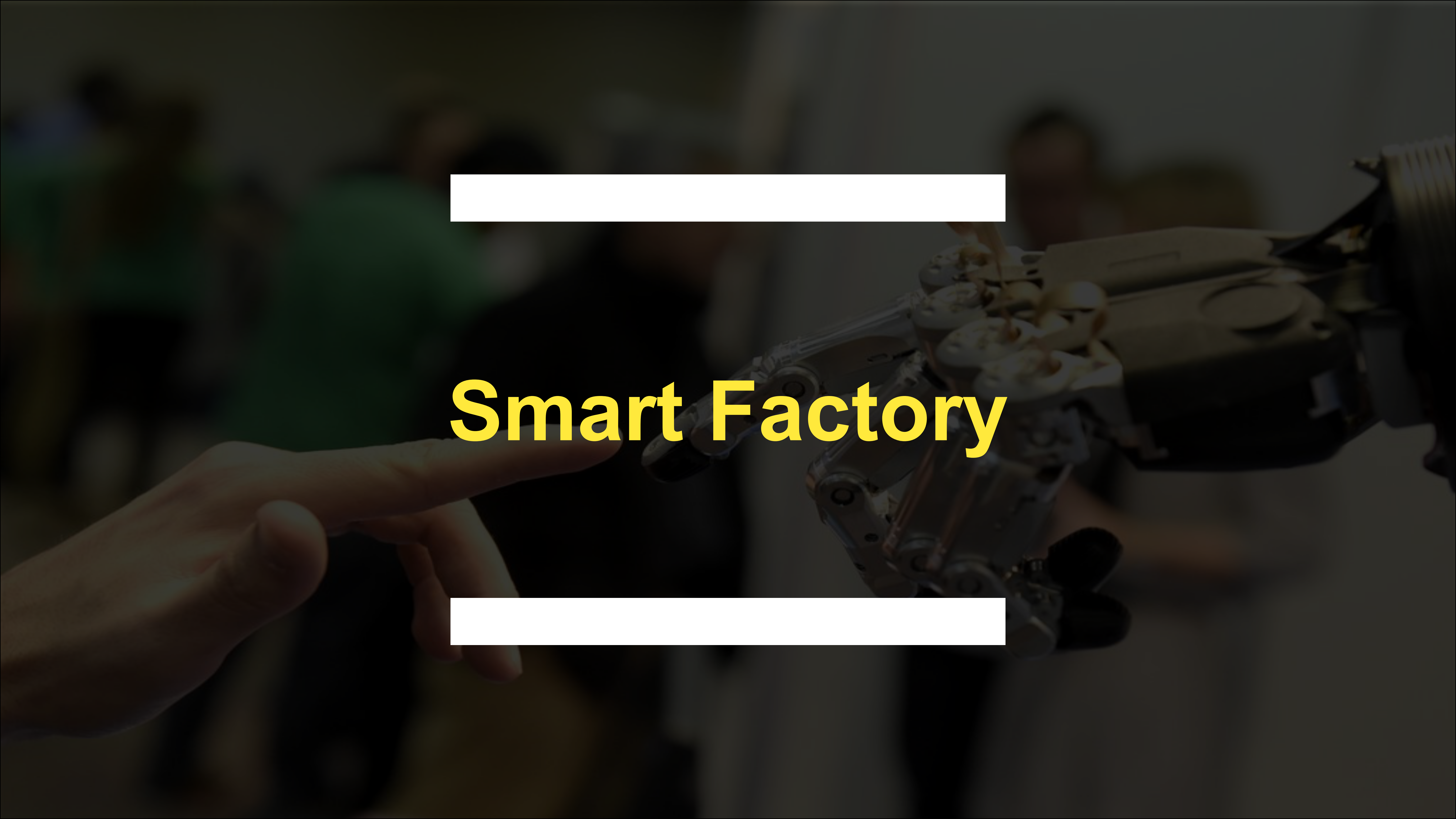


Industrial Data Science & Introduction to Deep Learning

Abbas Chokor, Ph.D.

AGENDA

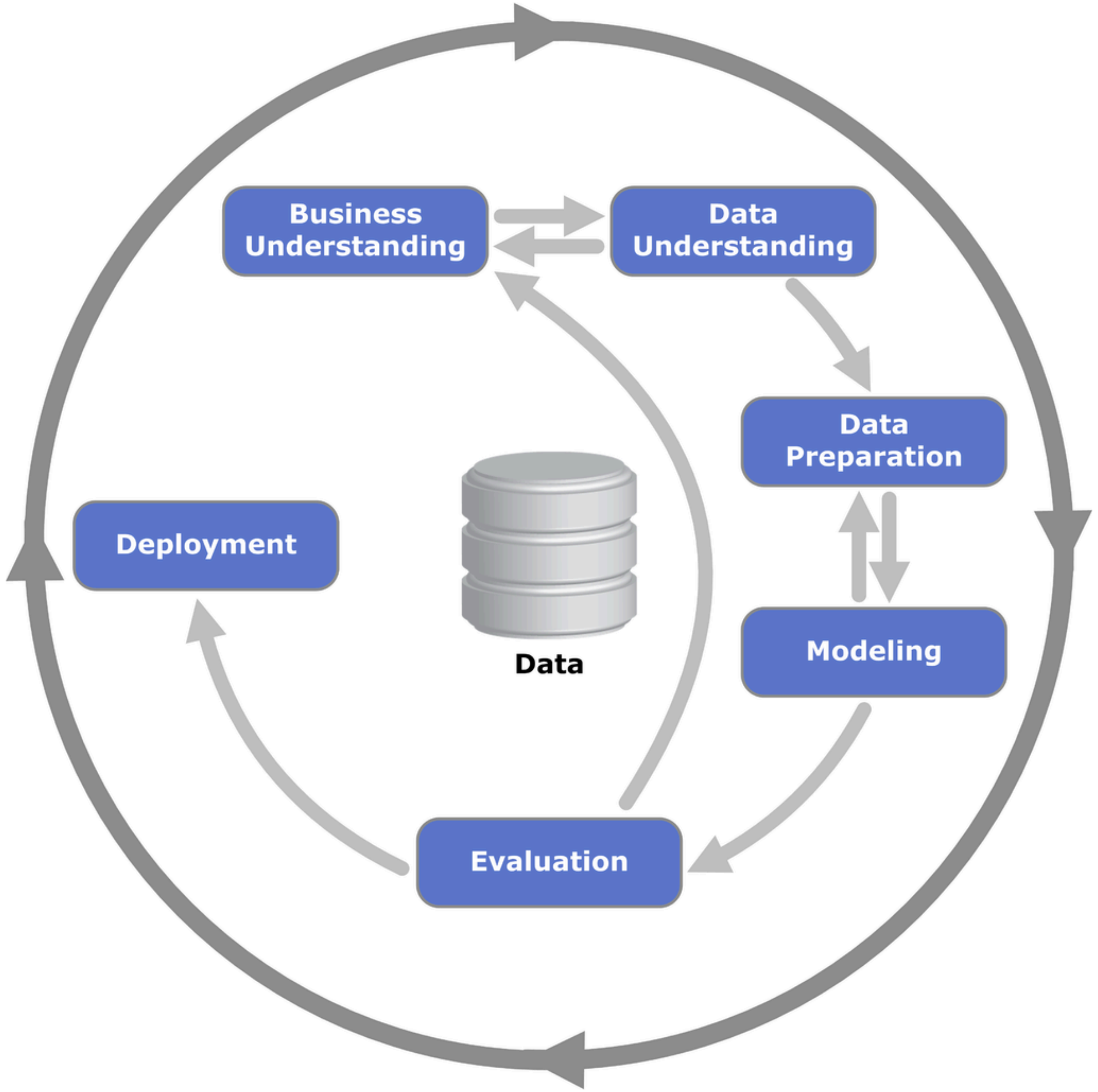
- Smart Factory
- PHM: Fault Diagnostics & Prognostics
- Example Diagnostics
- Example Prognostics



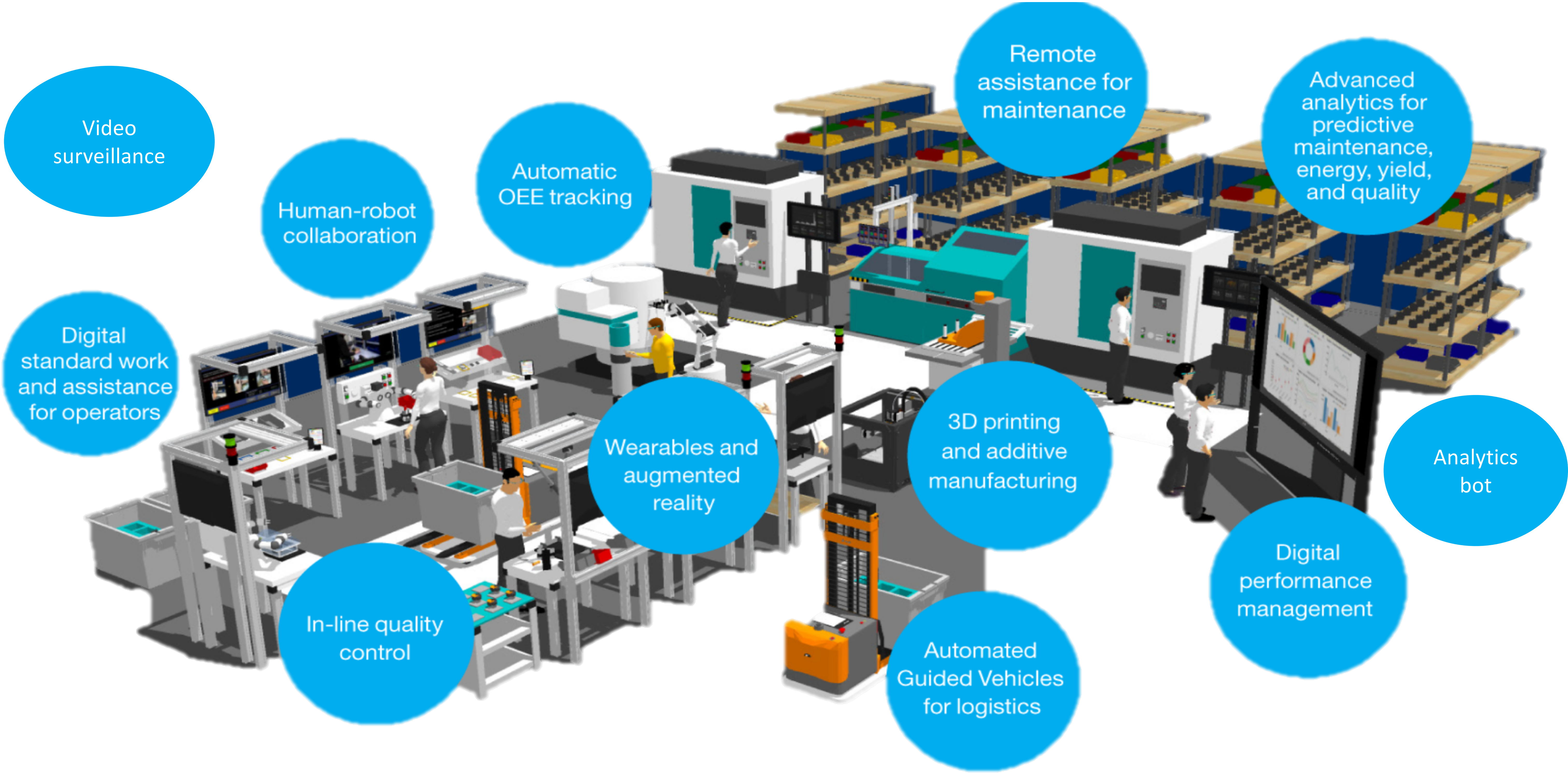
Smart Factory

Industrial Data Science

The results-oriented application of analytical techniques (from physics, statistics, machine learning, operations research, advanced visualization, etc.) to derive value from data.

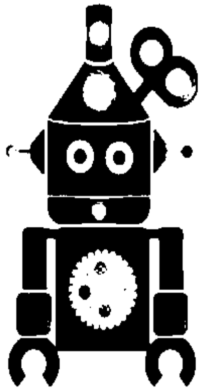


Smart Factory

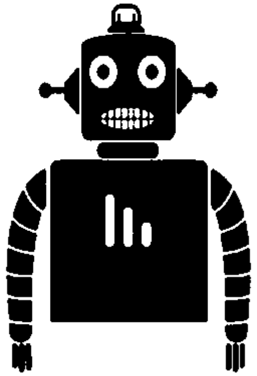


Smart Factory

Smart Factory 1.0
Rule Based



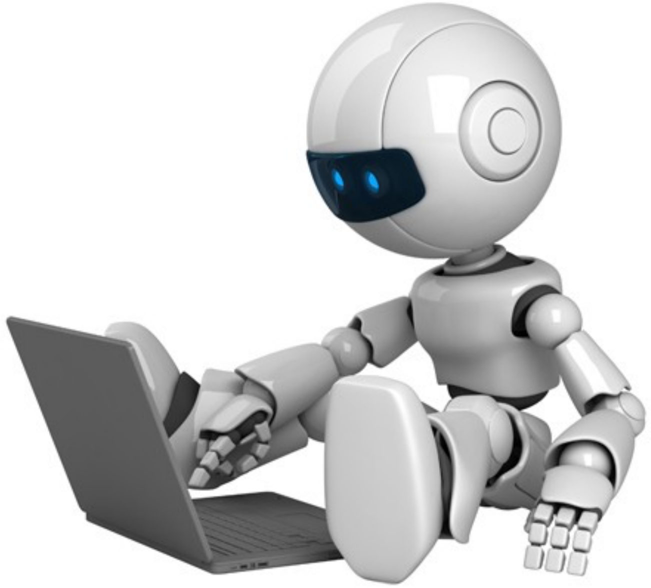
Smart Factory 2.0
Simple Machine Learning



Smart Factory 3.0
Deep Learning

What’s next?

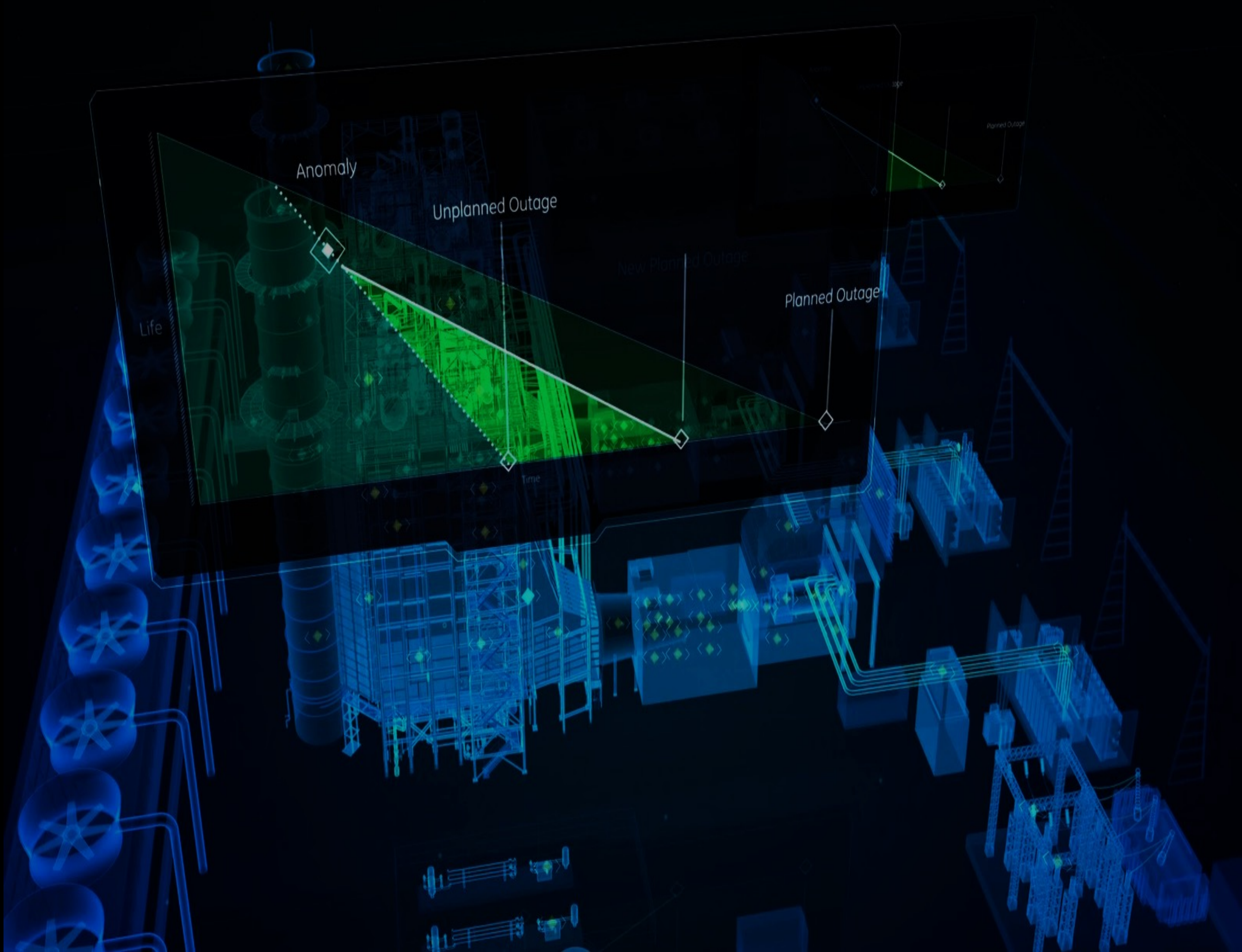
Smart Factory 4.0
Adaptive Learning



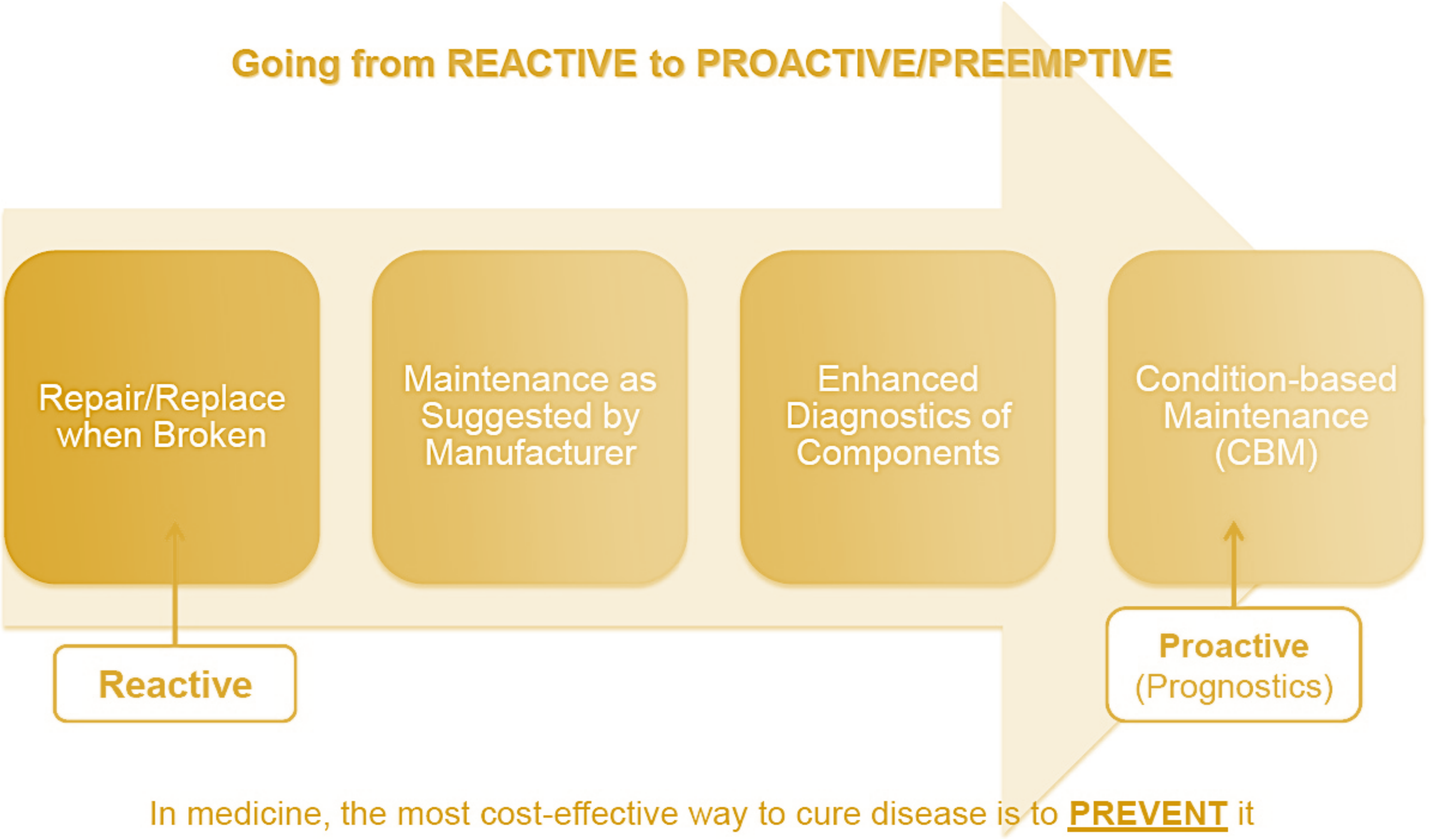
PHM
Fault Diagnostics & Prognostics

WHAT IS PHM?

A discipline that links studies of failure mechanisms to system lifecycle management through advanced data analytics.



WHY SHOULD WE MOVE TOWARDS PHM?



WHO IS ALREADY DOING PHM?



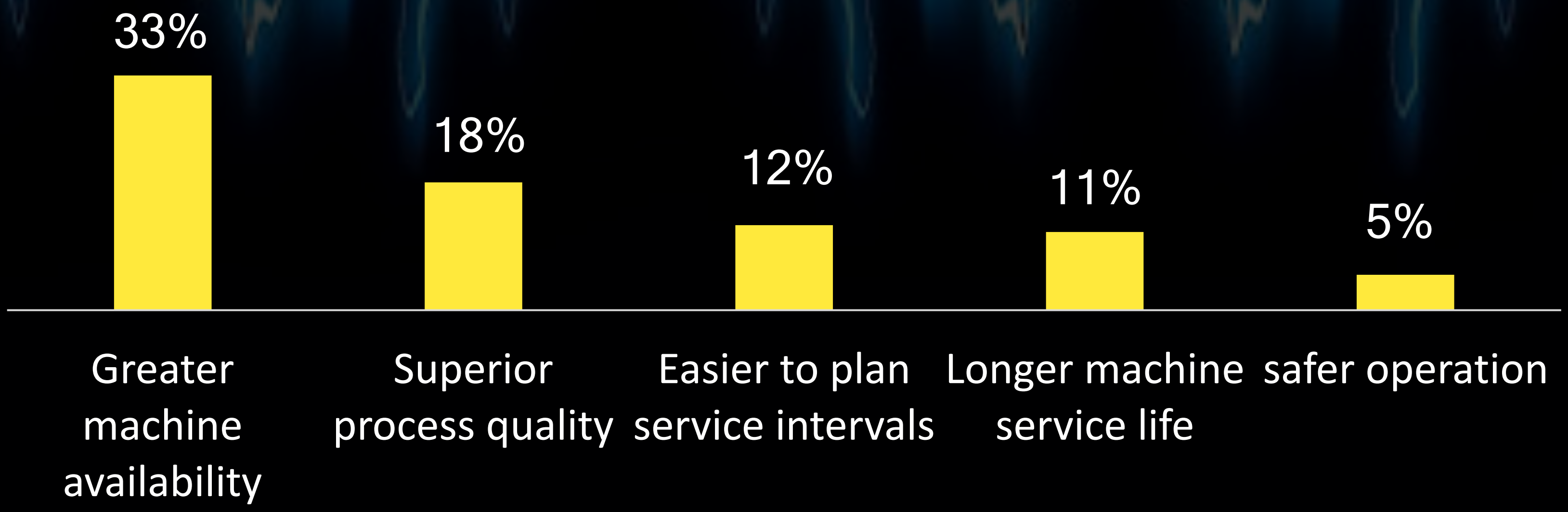





81% of manufacturers are tackling PHM **intensively…**

A recent survey* found that:

… achieving additional performance **gains**! **79%**



** Source: Roland Berger survey, January 31, 2017; n= 153*

PHM Initiatives

- Uniting groups in a PHM community
- Offering PHM 101
- Developing an overall PHM strategy to prioritize projects

PHM ROI potential is substantial:

Prevent Emergency Repairs/Breakdowns

- Full failure
- Spare Parts
- Labor
- Unfinished Products
- Contaminants Removal
- Late-delivery Penalties

Reduce Unnecessary Changes

- Parts Cost
- Labor
- Downtime
- Warranty

PHM IN A NUTSHELL

Establish the business opportunity

- STEP 1 Identify the target outcome
- STEP 2 Inventory data sources

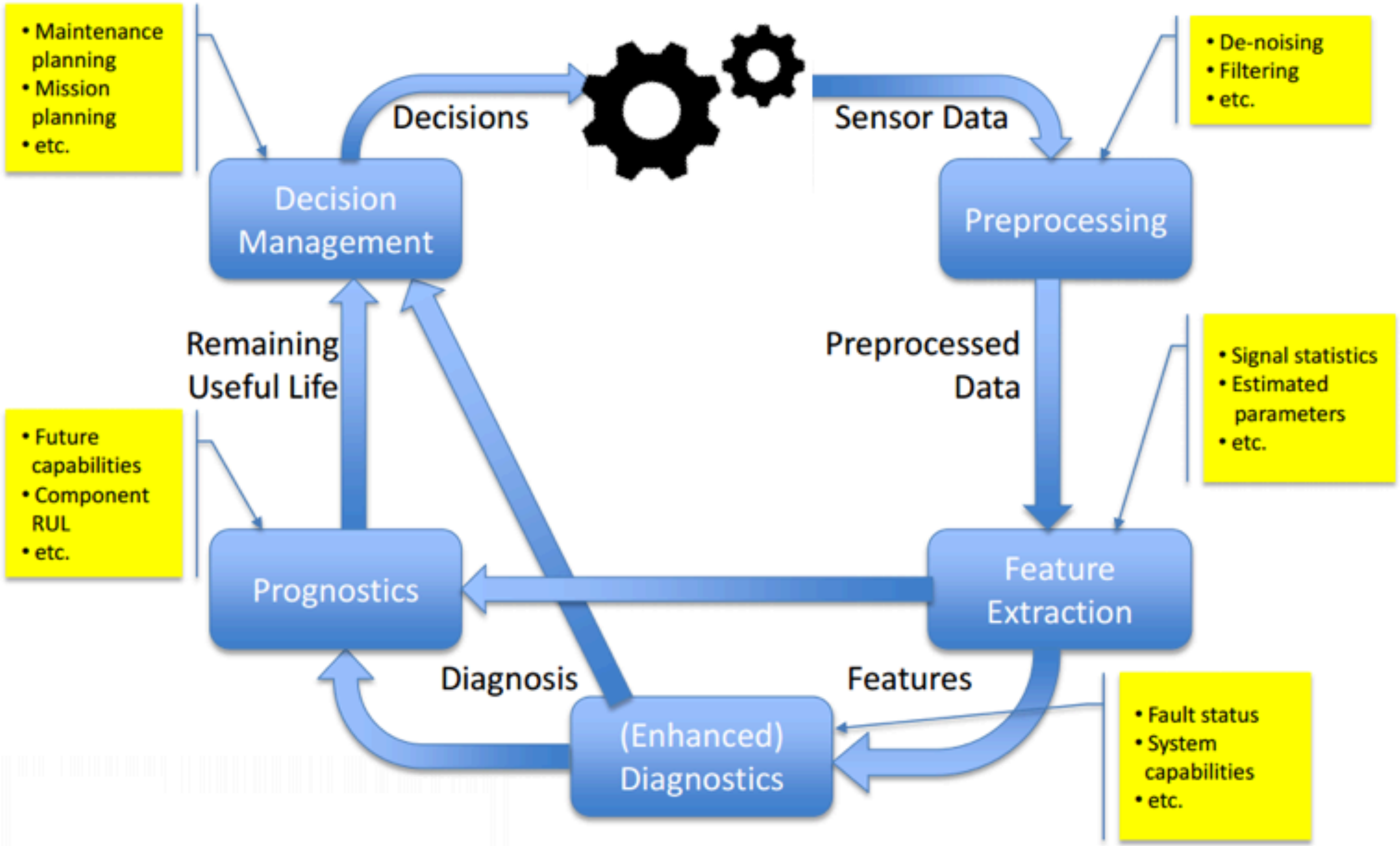
Gain predictive insight

- STEP 3 Capture and combine data
- STEP 4 Model, test, and iterate

Translate insight into action

- STEP 5 Validate model in a live operational setting
- STEP 6 Integrate into operations

PHM SYSTEM ARCHITECTURE



WHERE TO START?

FMEA / FMECA

- Failure modes
- Effects (and Criticality) – which failure modes to go after

Fault Tree Analysis

- Propagation Models

Designers / Reliability Engineers

- System knowledge and insight
- Expected / nominal behavior of the system

Failure Testing / Accelerated Life Testing

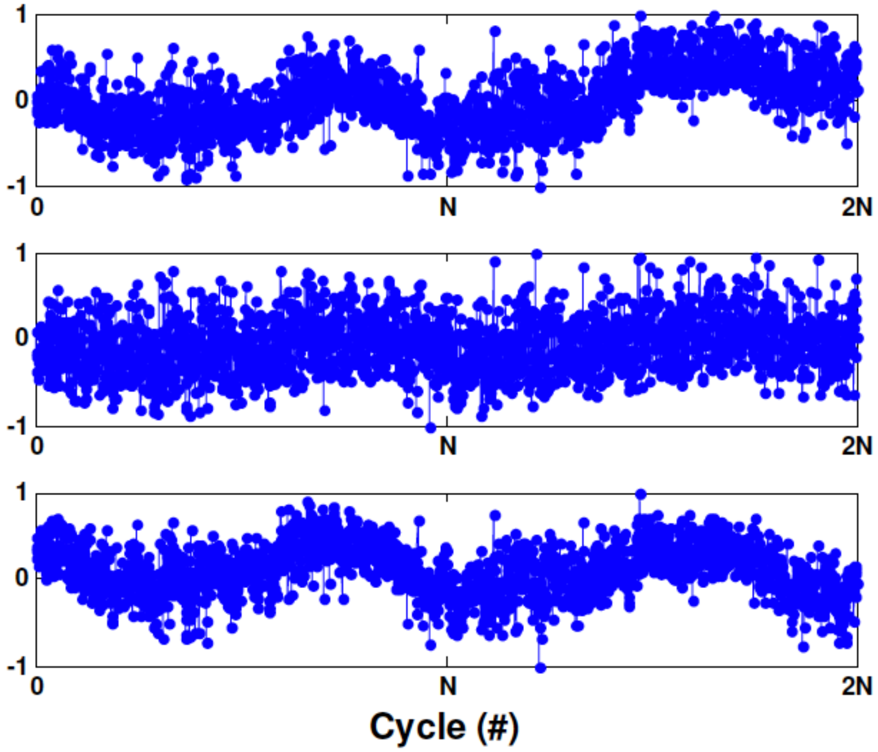
- Data
- Failure signatures
- Effects of environmental conditions

Fielded Systems

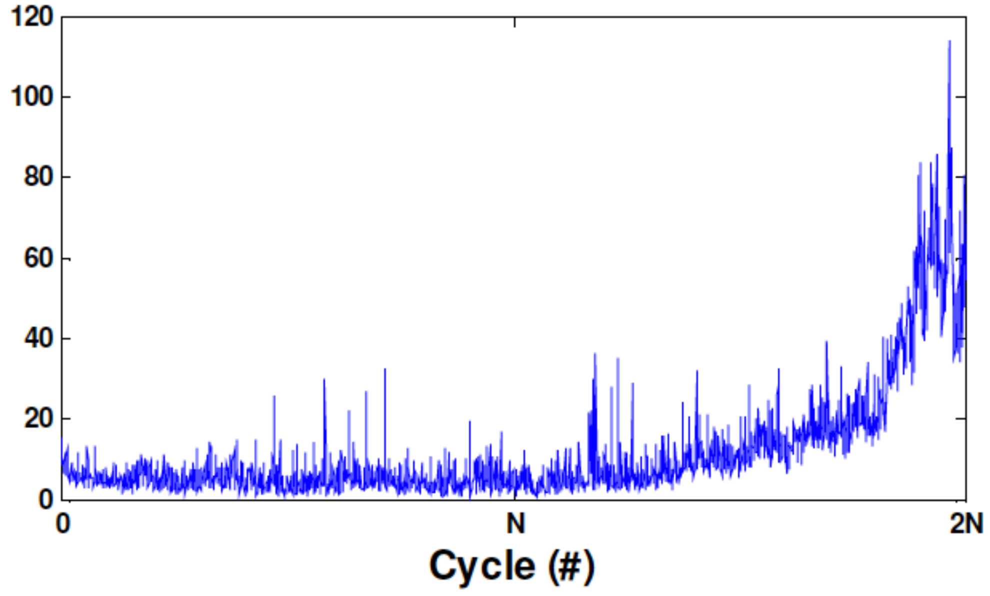
- Sensors measurements
- Machine Statistics
- Maintenance logs
- Performance Validation

PHM ANALYSIS: PREPROCESSING

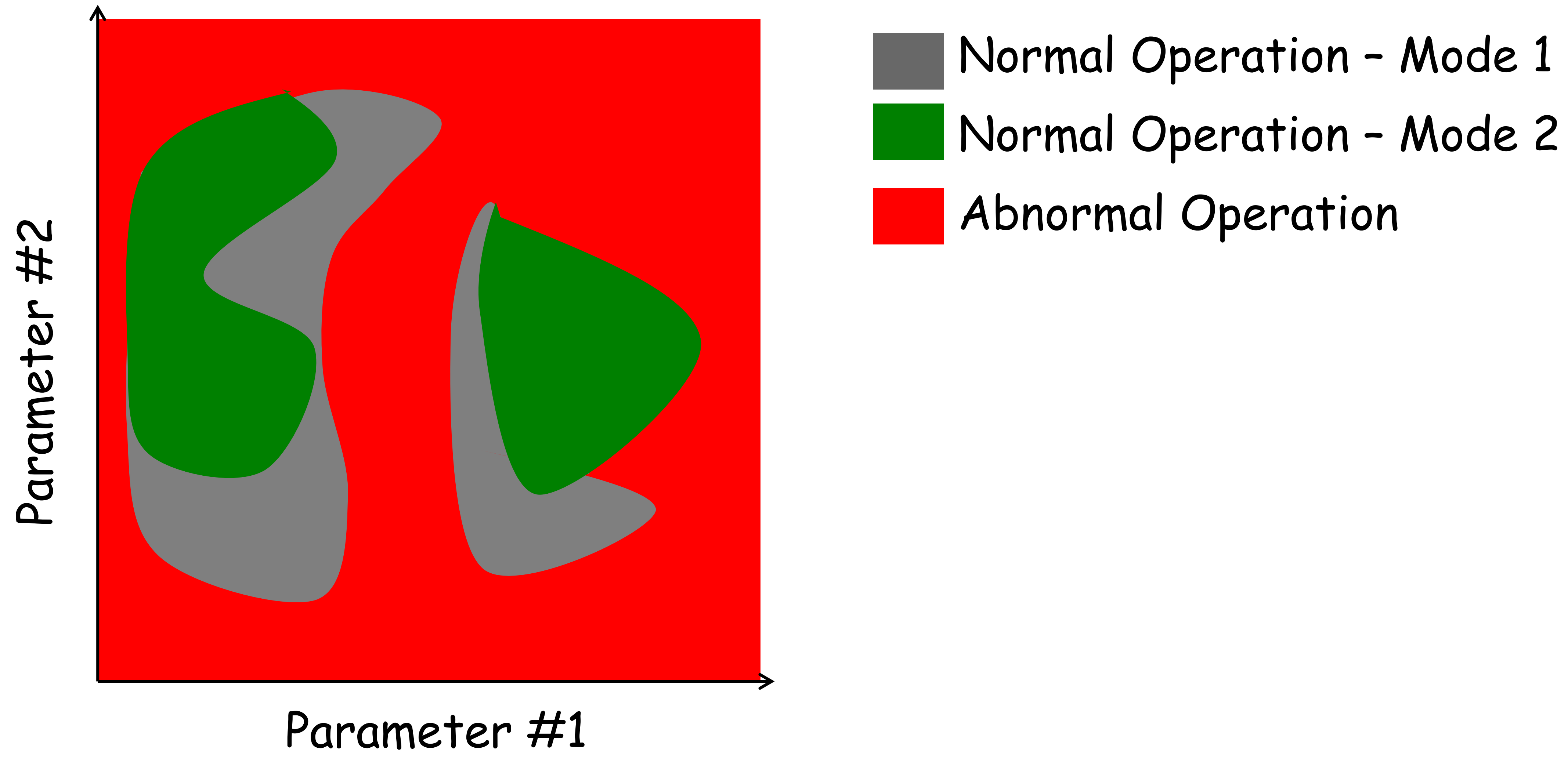
Raw Measurements



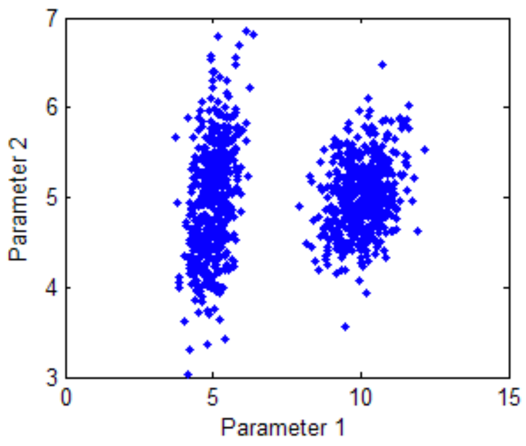
Features



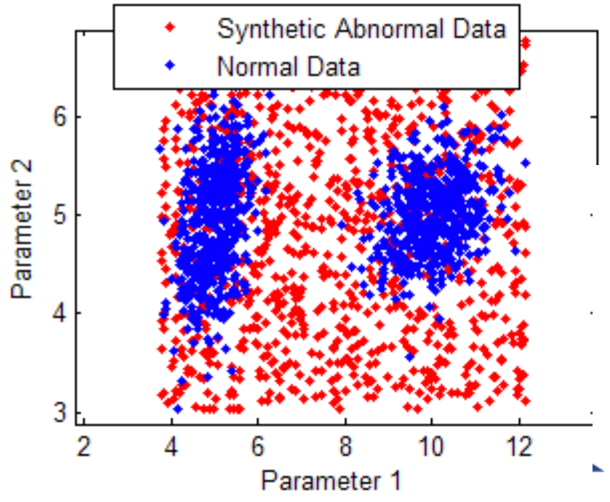
PHM ANALYSIS: FAULT DETECTION



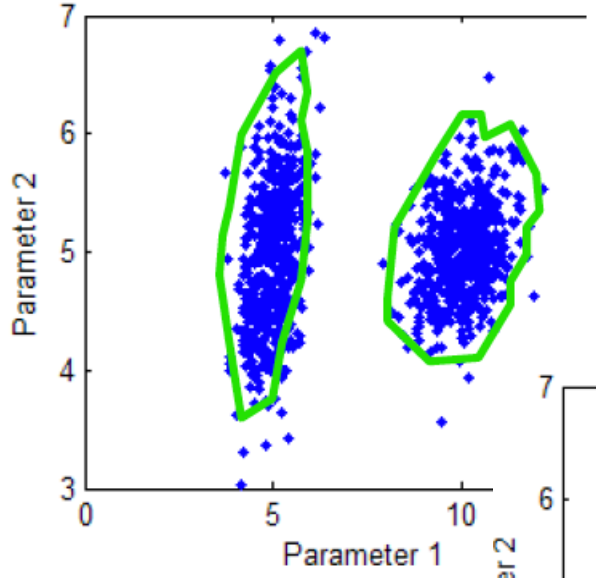
PHM ANALYSIS: BUILD A CLASSIFIER



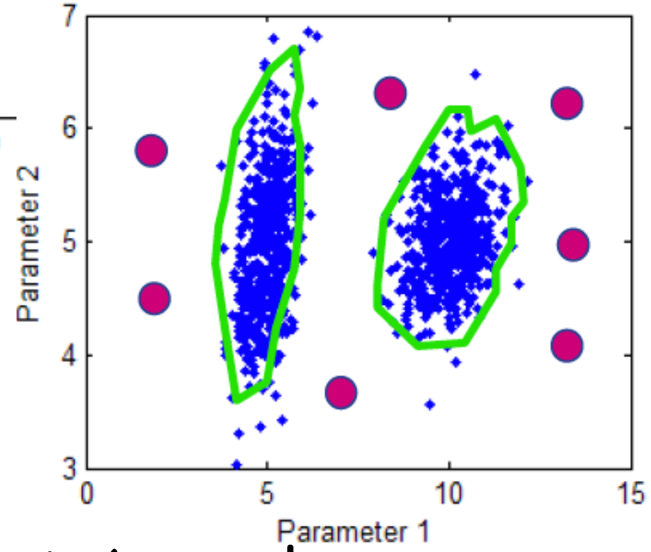
Normal Data



Generate Abnormal Data

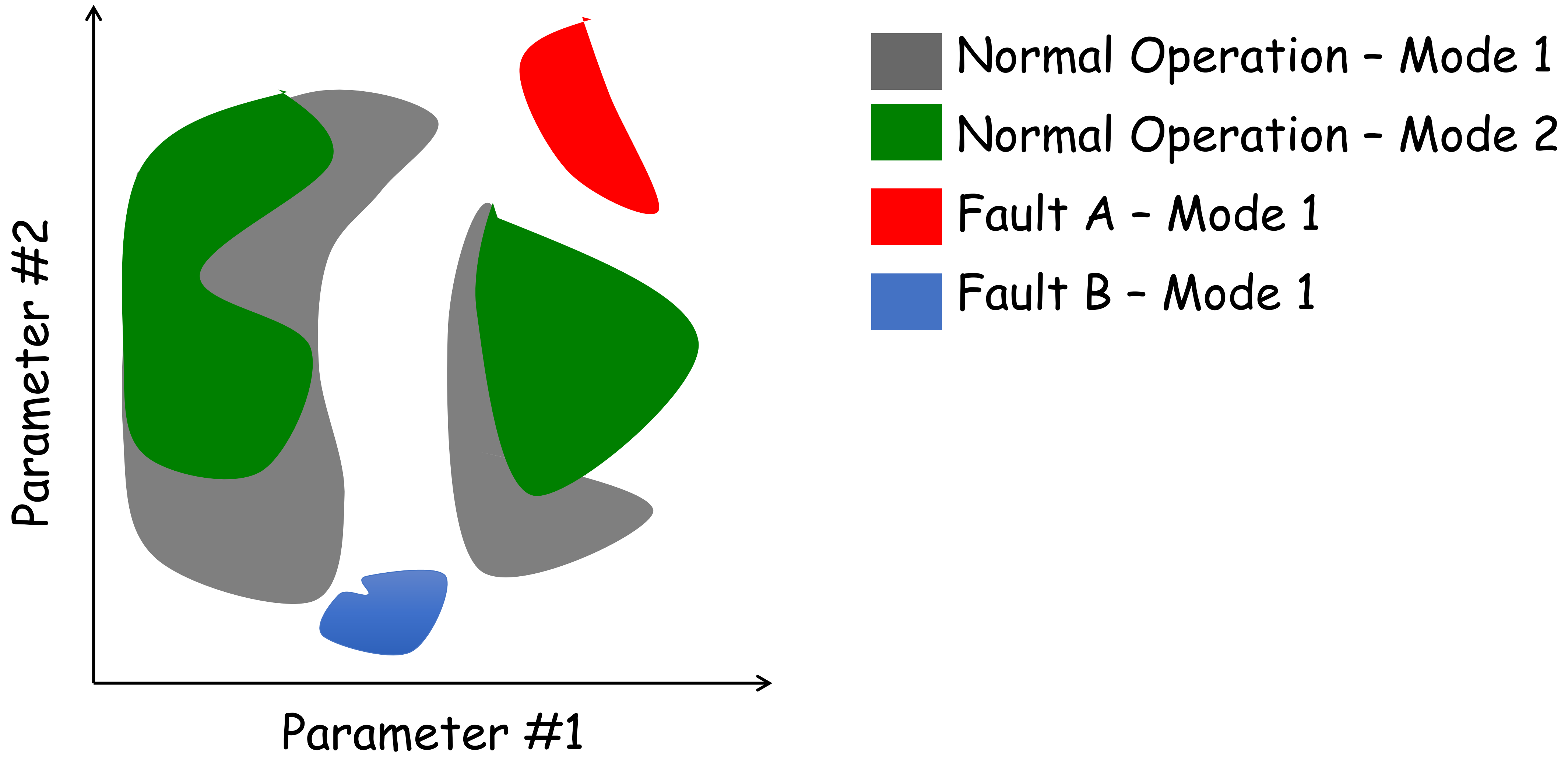


Build a RF Classifier

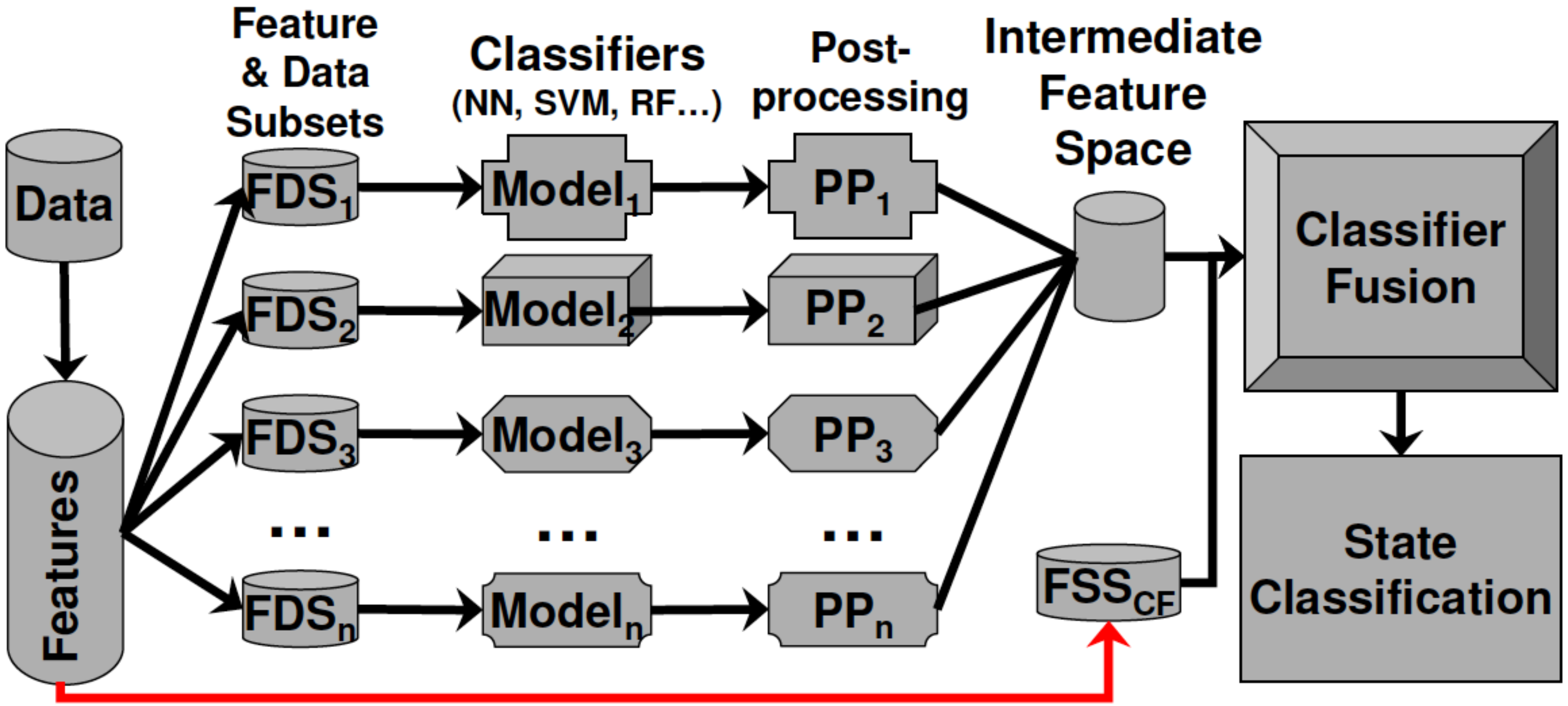


Detect Anomaly

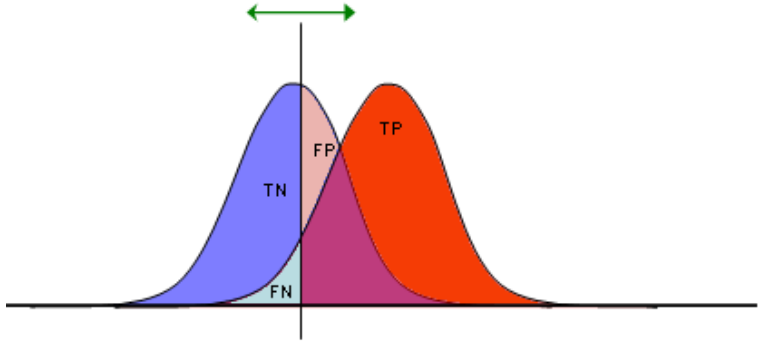
PHM ANALYSIS: FAULT DIAGNOSTICS



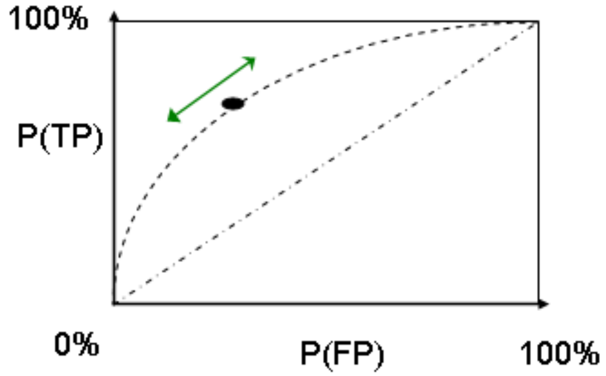
PHM ANALYSIS: FAULT DIAGNOSTICS

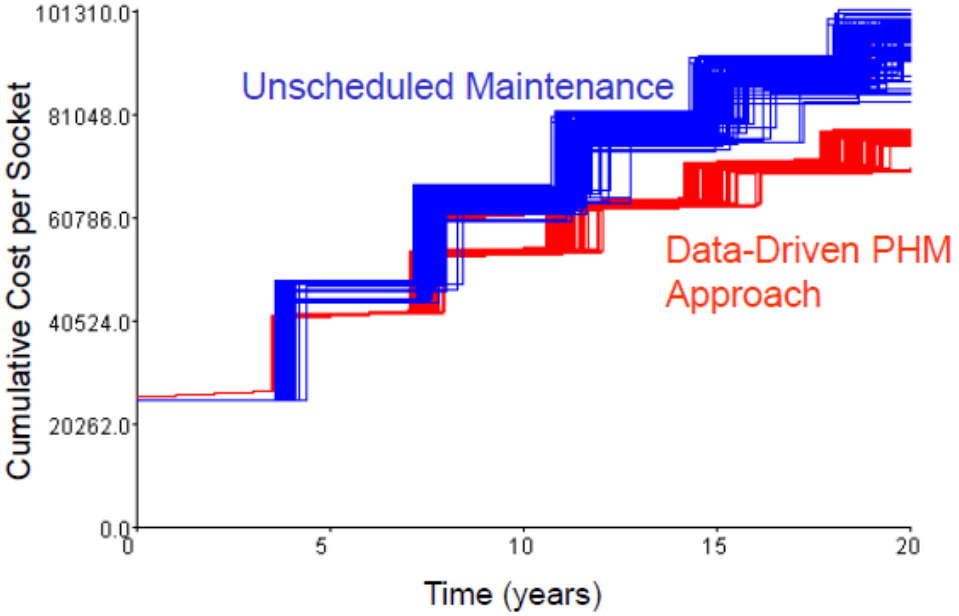


PHM ANALYSIS: MODEL EVALUATION



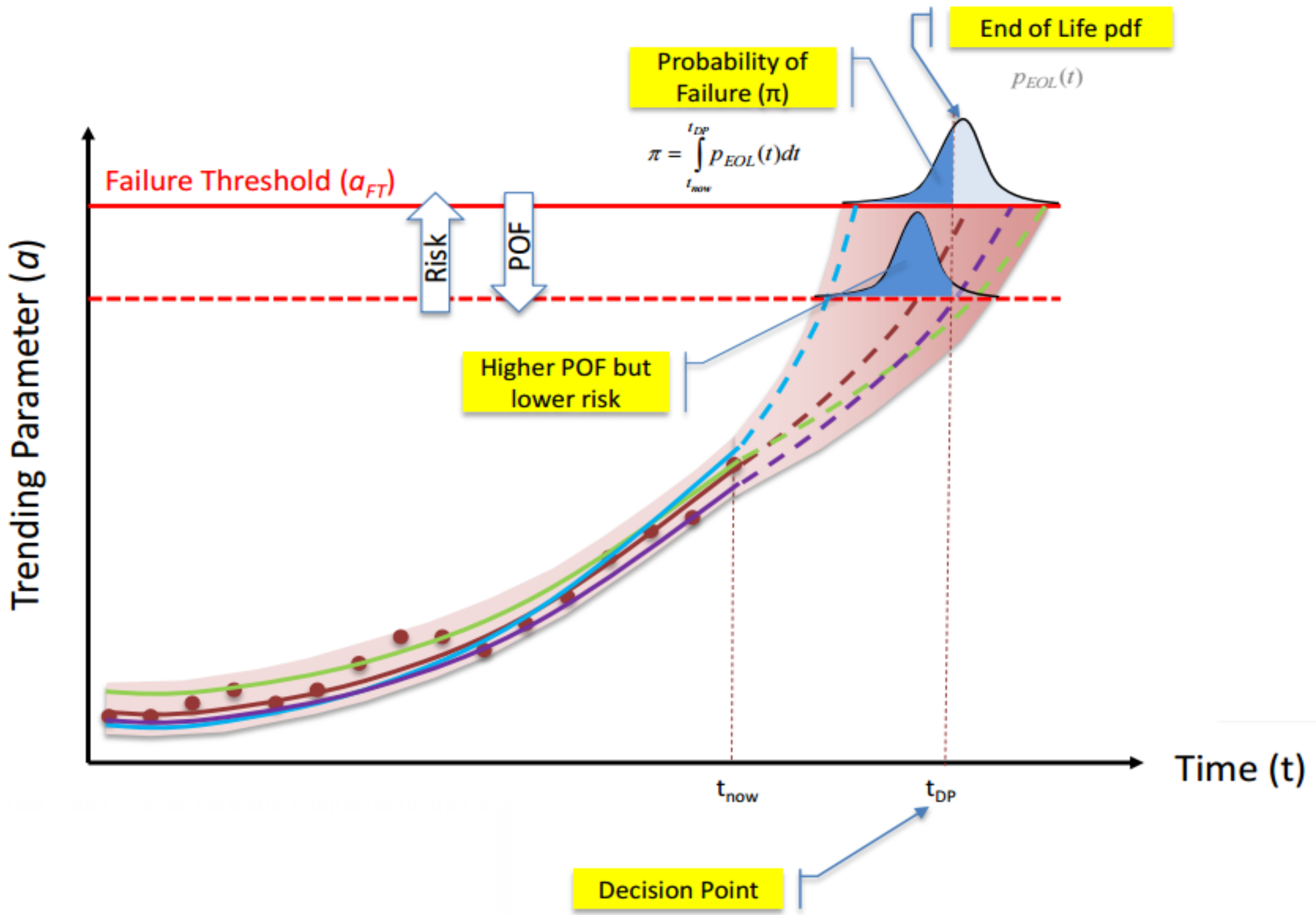
TP	FP
FN	TN
1	1





PHM ANALYSIS: PROGNOSTICS

PHM ANALYSIS: PROGNOSTICS



PHM ROI

- ROI relative to unscheduled maintenance gives

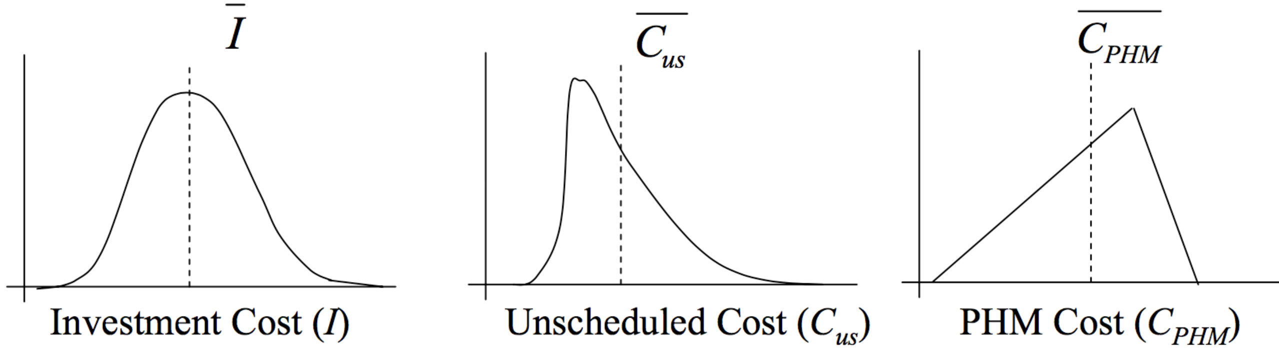
$$ROI = \frac{(C_{us} - I_{us}) - (C_{PHM} - I_{PHM})}{(I_{PHM} - I_{us})} - 1$$

where,

C_{us} = total life cycle cost using unscheduled maintenance
C_{PHM} = total life cycle cost using the selected PHM approach
I_{us} = unscheduled maintenance investment cost
I_{PHM} = PHM investment cost

- By definition, $I_{us} = 0$ (contains no investment in PHM)

$$ROI = \frac{C_{us} - (C_{PHM} - I_{PHM})}{I_{PHM}} - 1$$

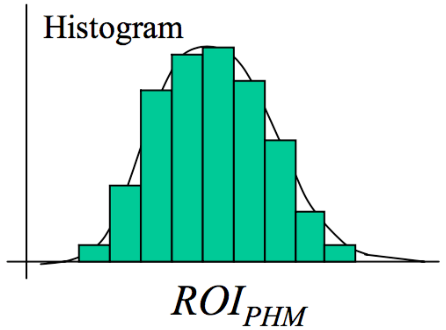


- Investment cost

$$I_{PHM} = C_{NRE} + C_{REC} + C_{INF}$$

where,

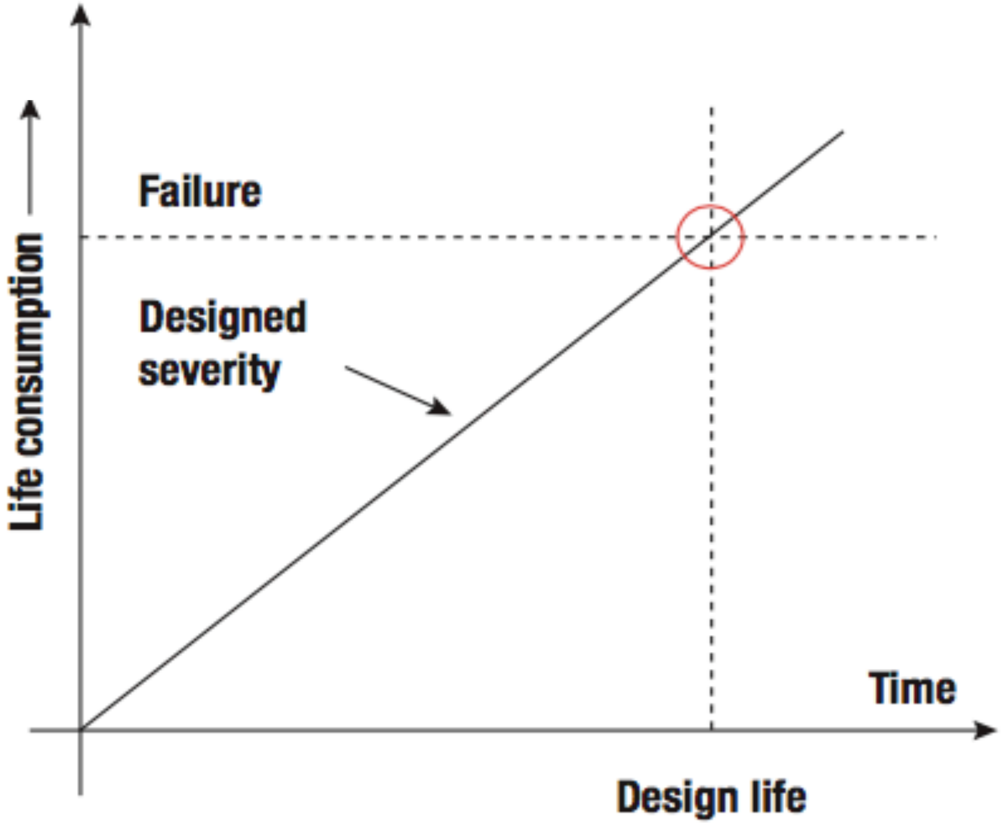
C_{NRE} = PHM non-recurring costs
C_{REC} = PHM recurring costs
C_{INF} = PHM infrastructure costs

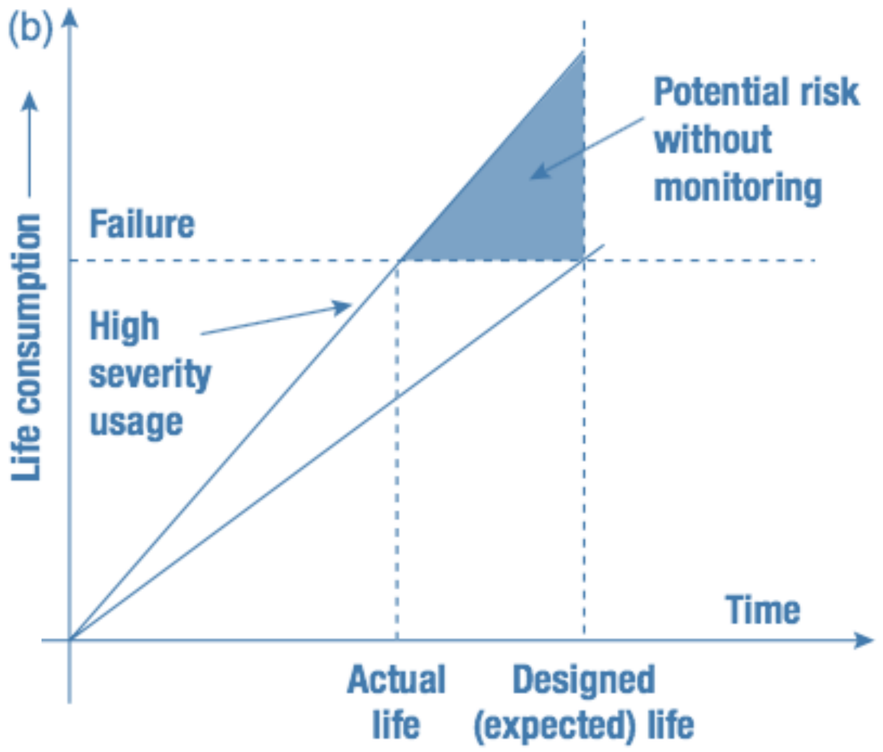


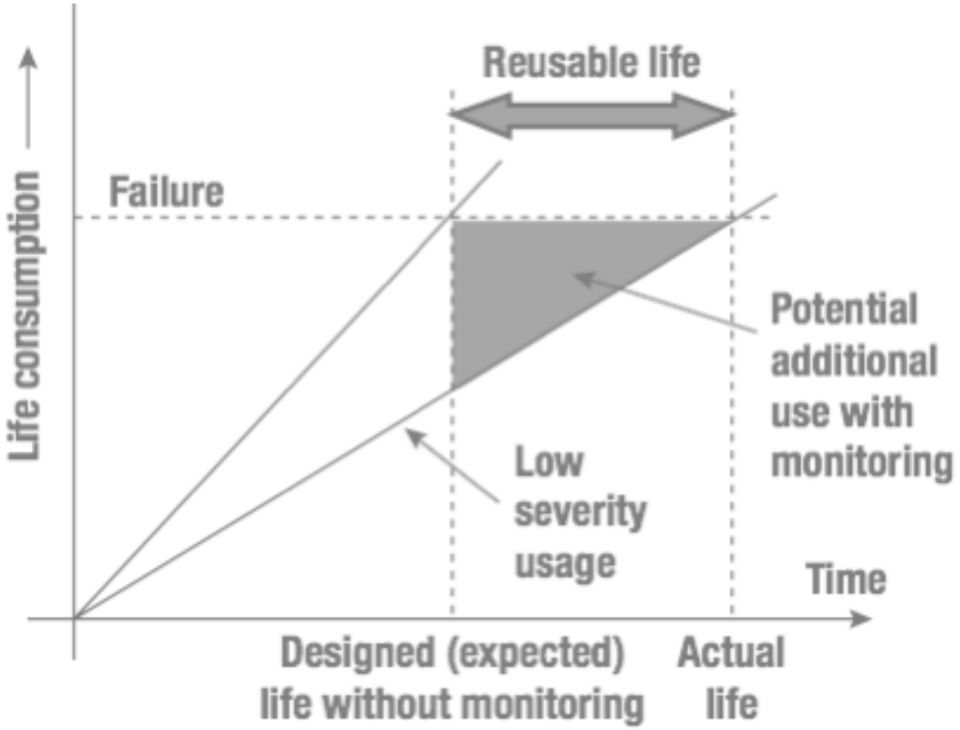
Value:

- Mean ROI
- ROI uncertainty
- ROI confidence

PHM ROI


Life consumption
Failure
Designed
severity
Time
Design life


(b)
Life consumption
Failure
Potential risk
without
monitoring
High
severity
usage
Time
Actual
life
Designed
(expected) life

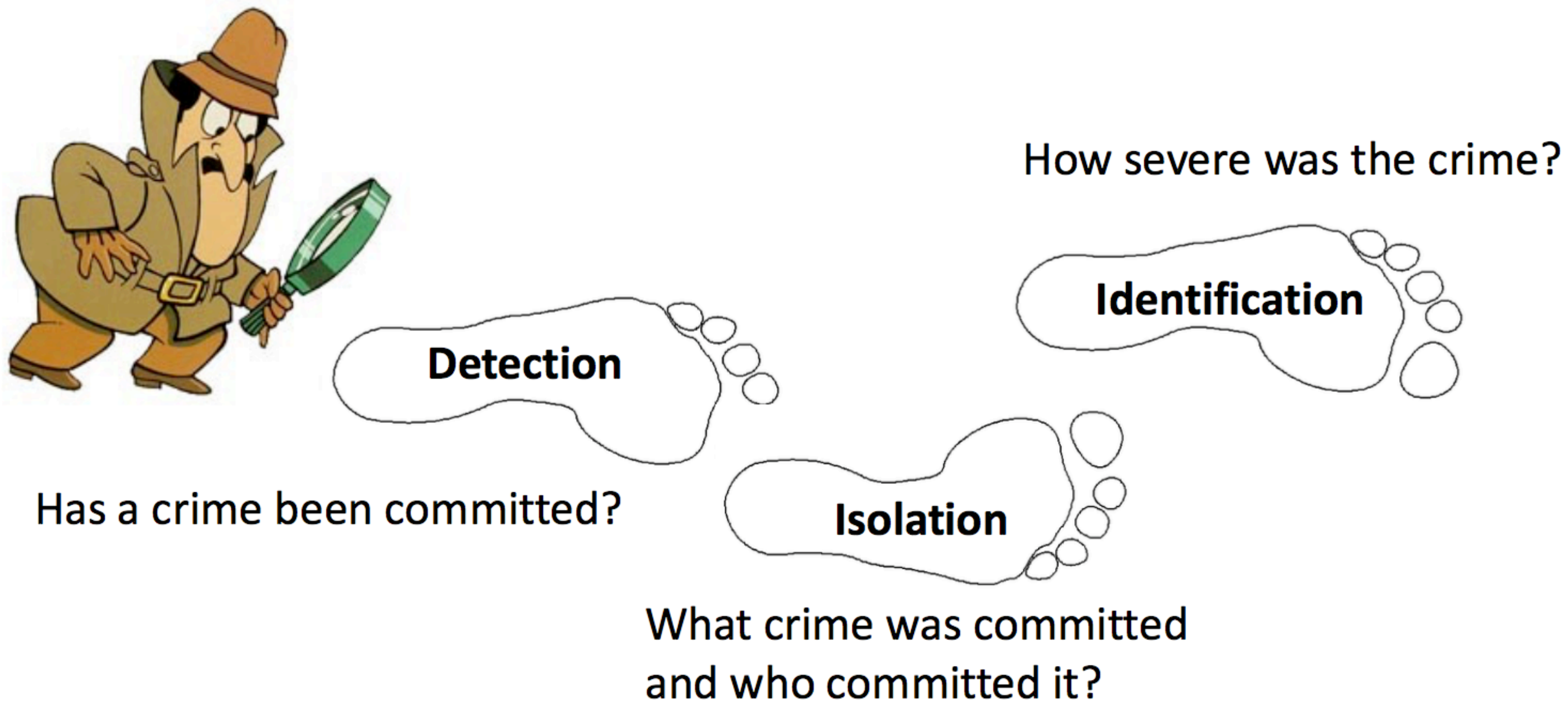

Reusable life
Life consumption
Failure
Potential
additional
use with
monitoring
Low
severity
usage
Time
Designed (expected)
life without monitoring
Actual
life

FAULT DIAGNOSTICS

WHAT IS DIAGNOSTICS?

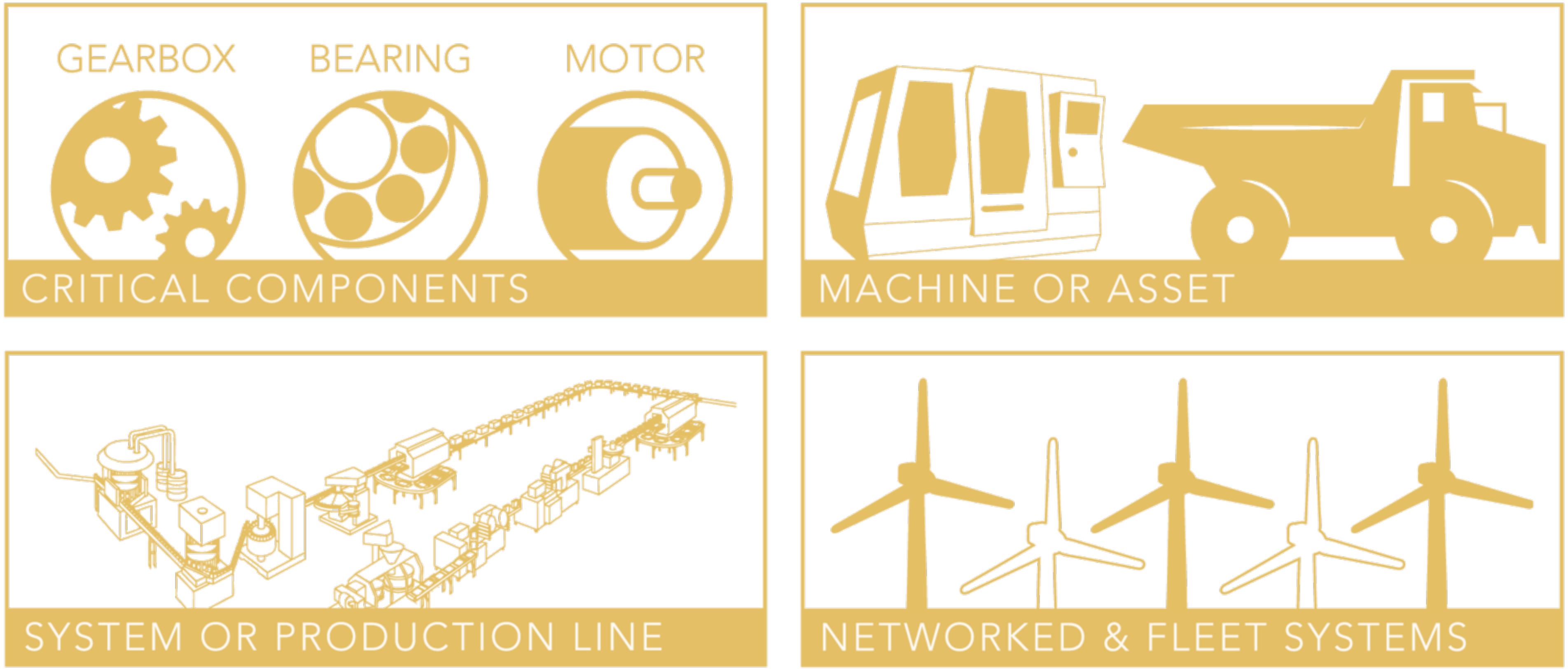
Diagnostics is the identification of the nature and cause of a certain phenomenon.

1. Fault Detection: An abnormal event reported.
2. Fault Isolation – Type, location and time of a fault.
3. Fault Identification – Size of the fault (severity)

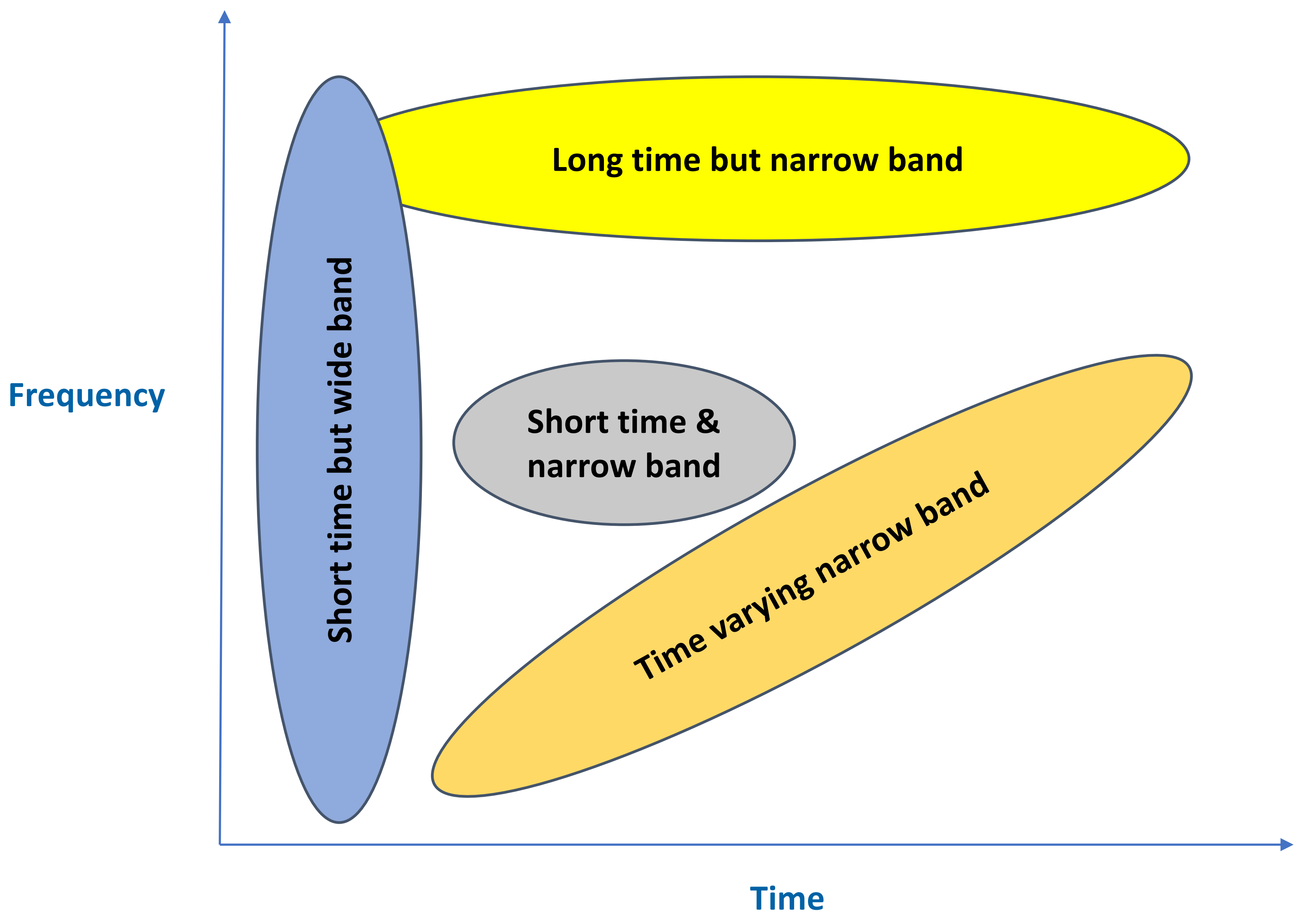


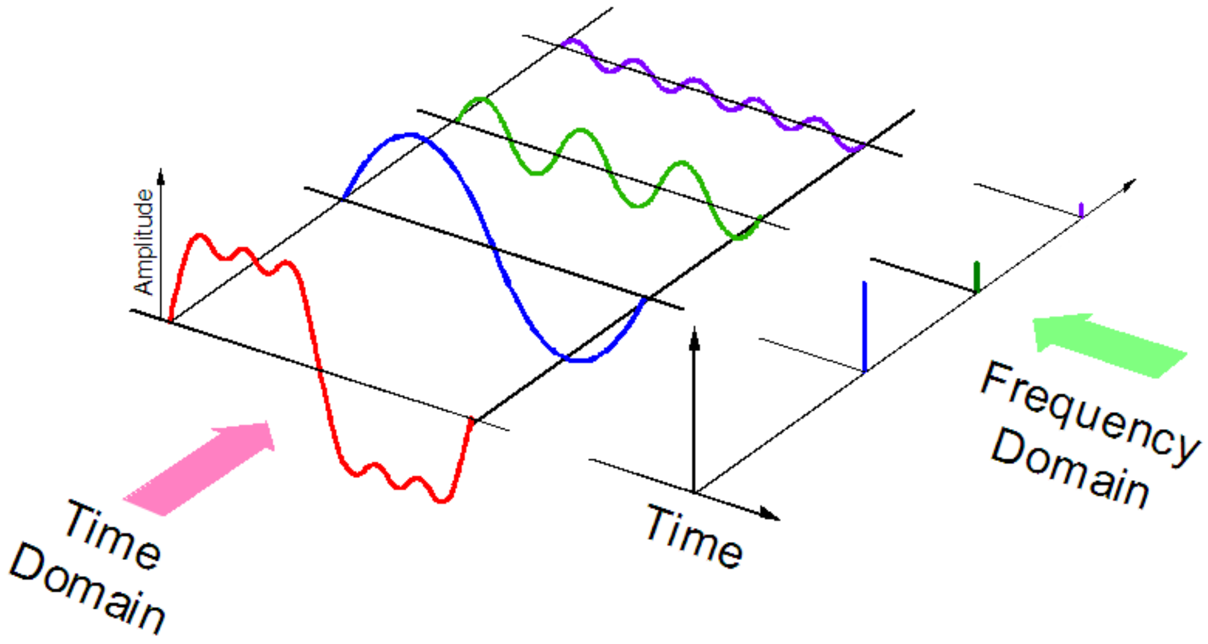
WHAT IS DIAGNOSTICS?

Pre-Processing → Feature Extraction → Classification → Reasoning



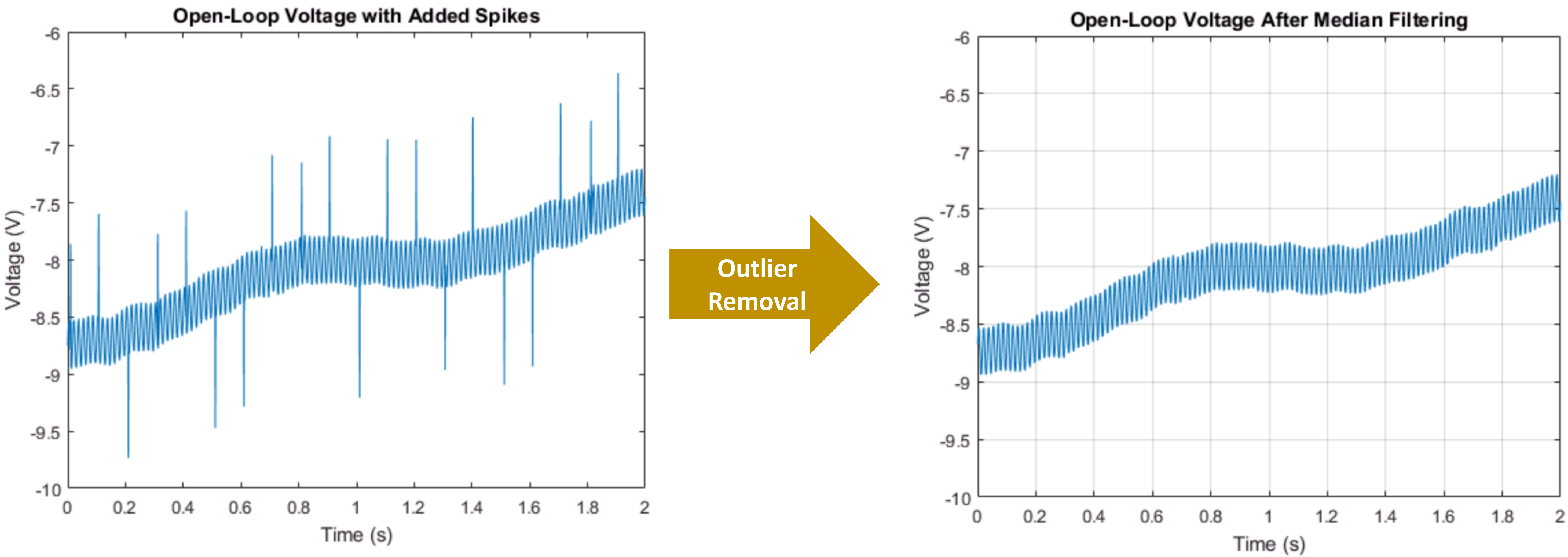
SIGNAL DATA: PREPROCESSING AND EXTRACTING FEATURES



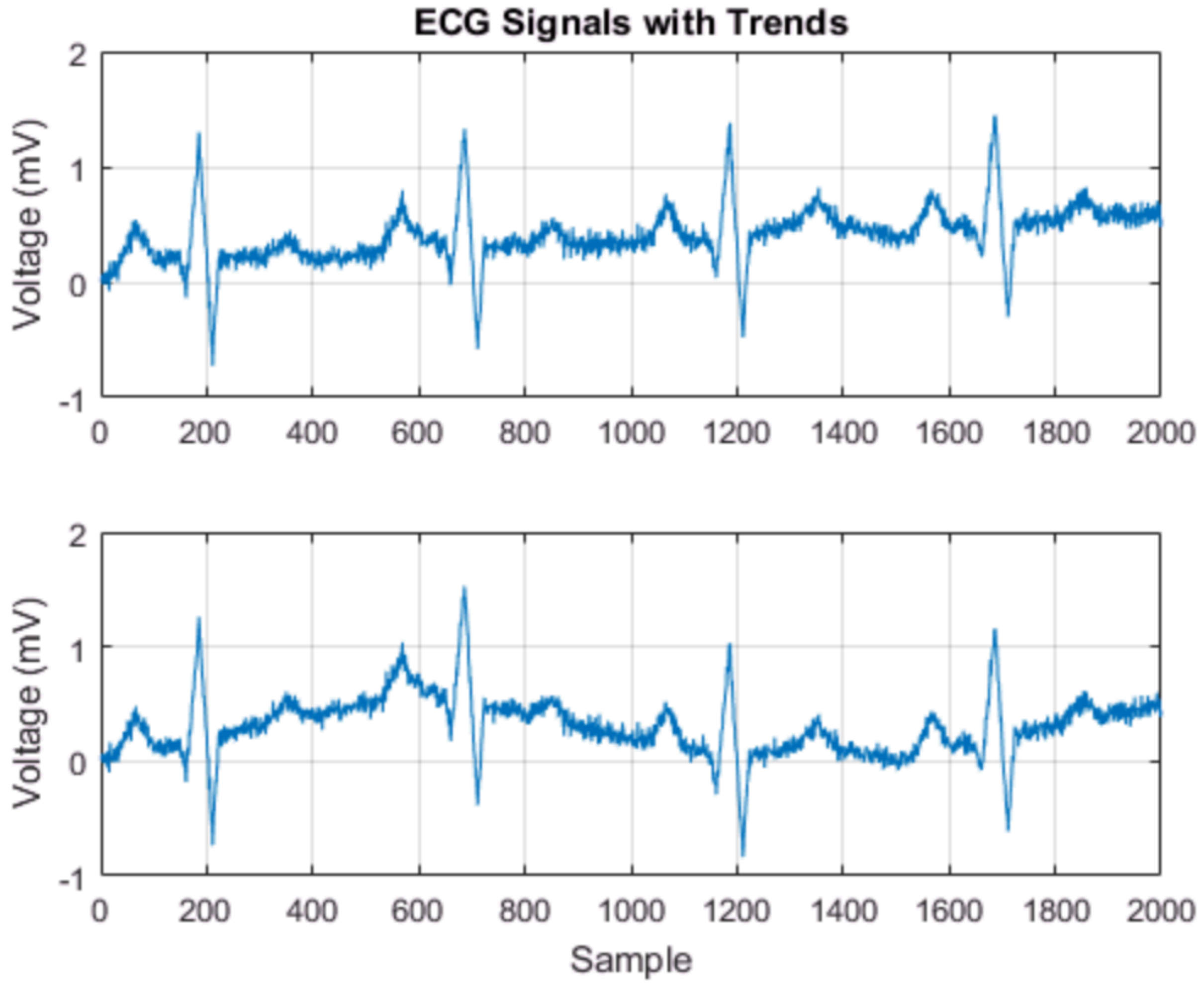


The Fourier Transform relates the function's time domain, shown in red, to the function's frequency domain, shown in blue.

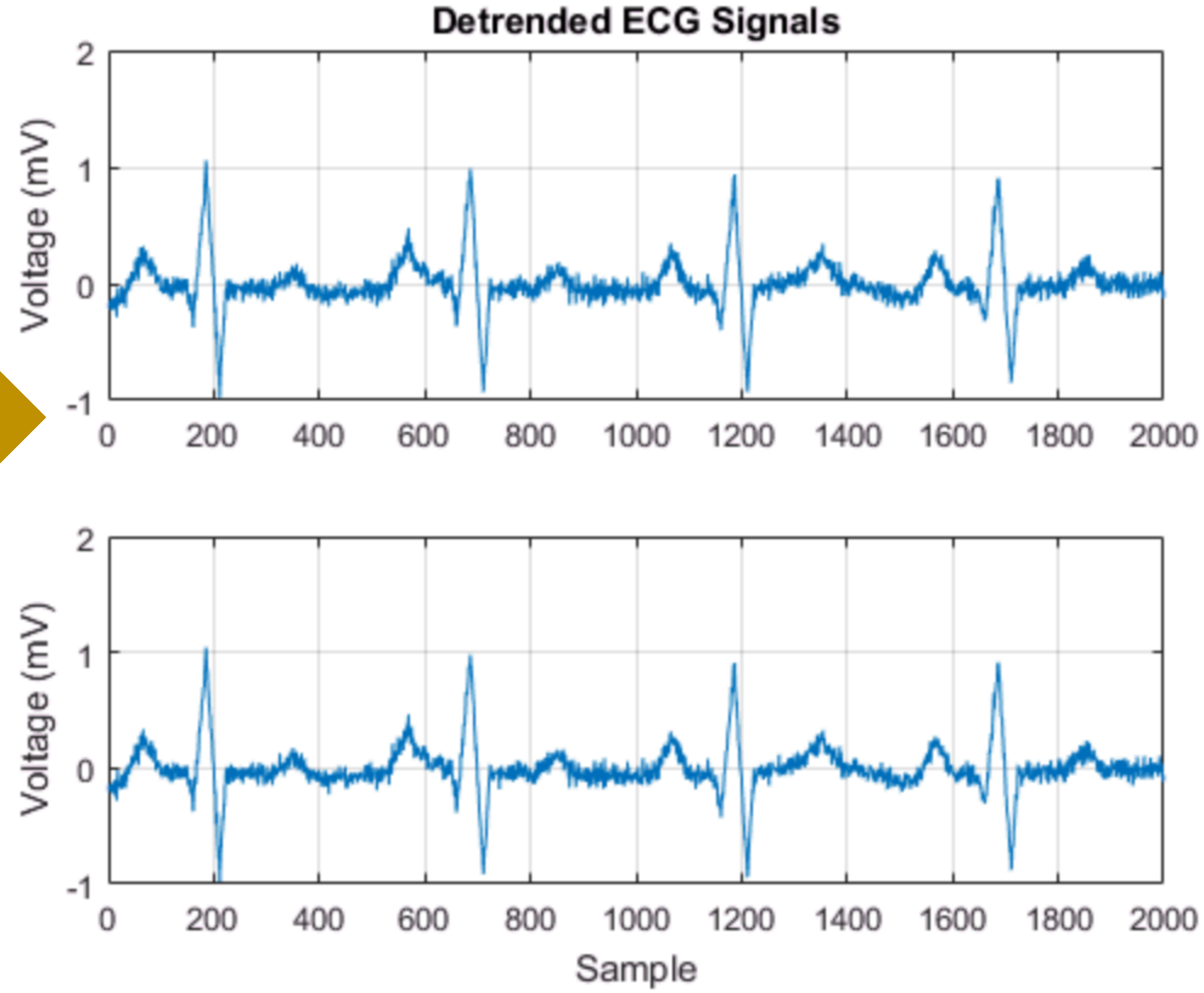
SIGNAL DATA: PREPROCESSING AND EXTRACTING FEATURES



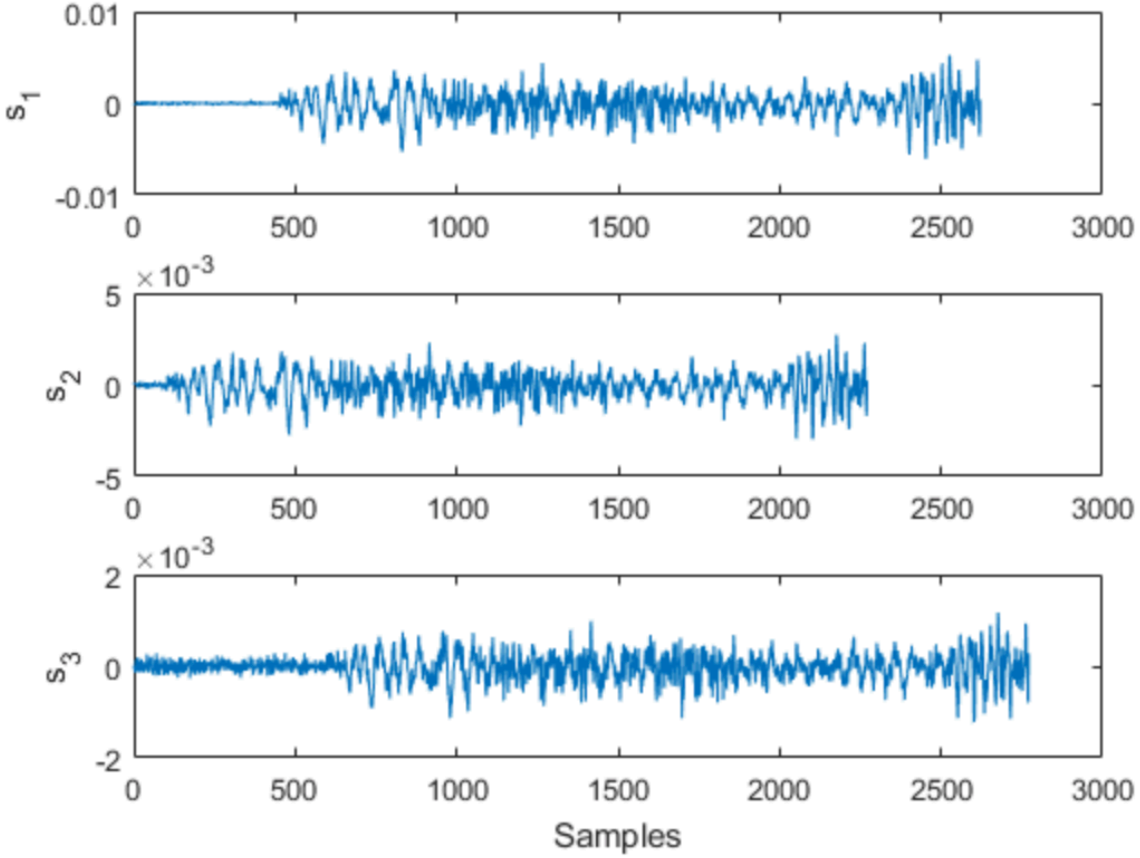
SIGNAL DATA: PREPROCESSING AND EXTRACTING FEATURES



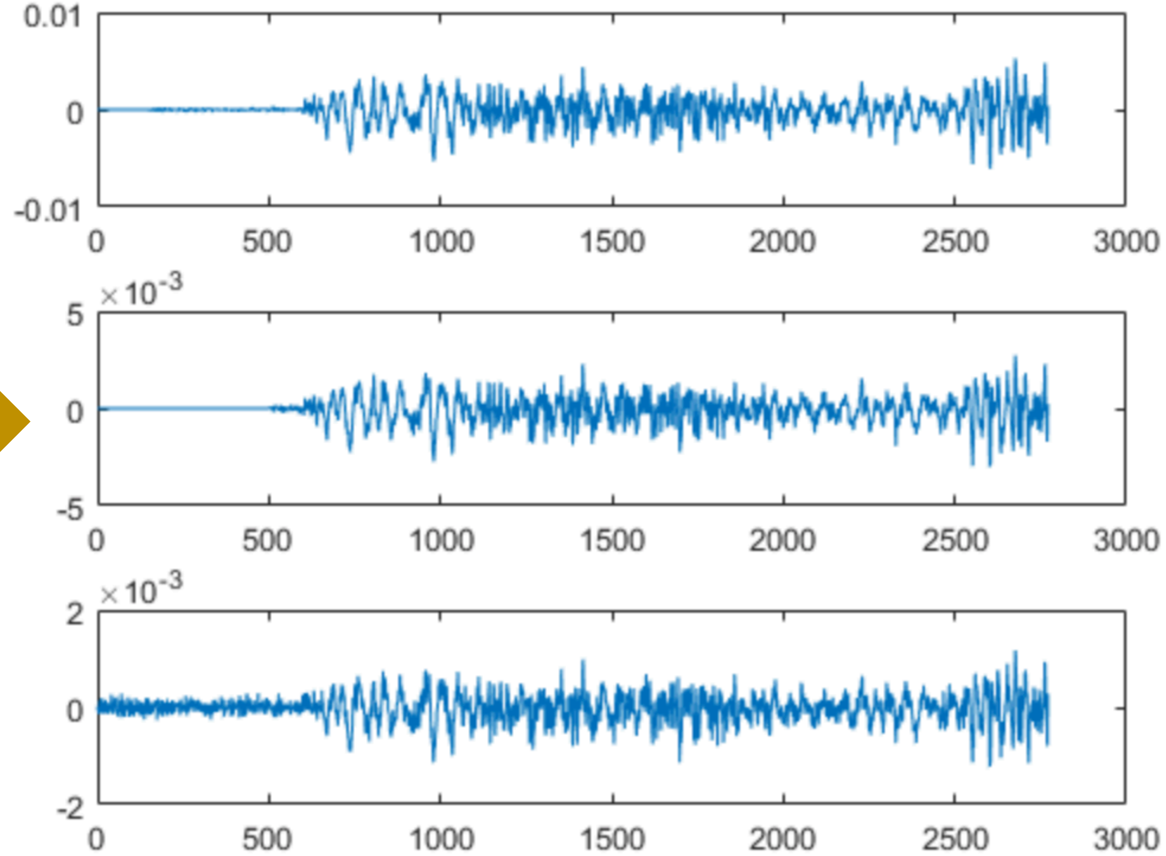
Detrending & Normalizing



SIGNAL DATA: PREPROCESSING AND EXTRACTING FEATURES

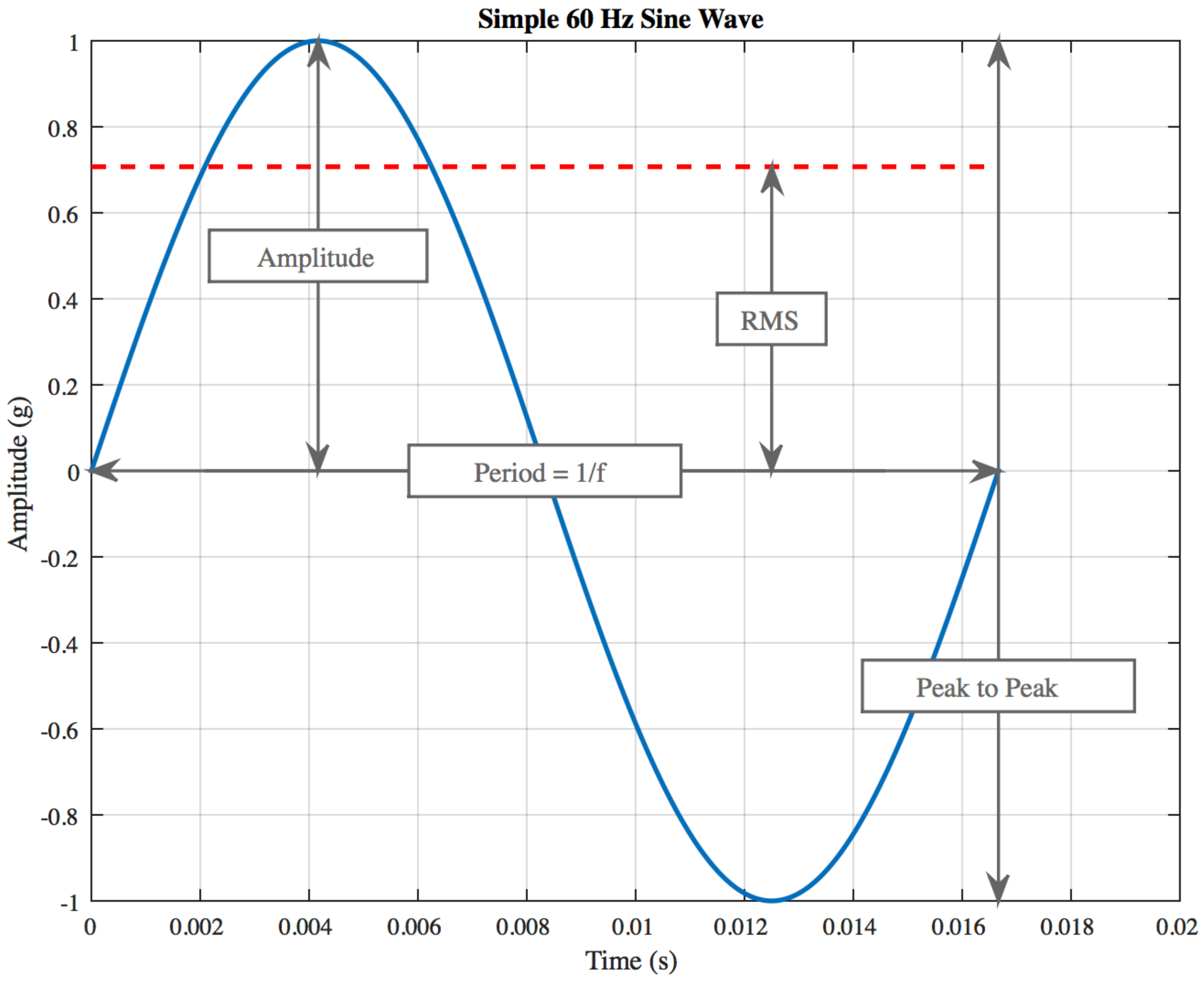


Synchronizing Signals with Different Start Times



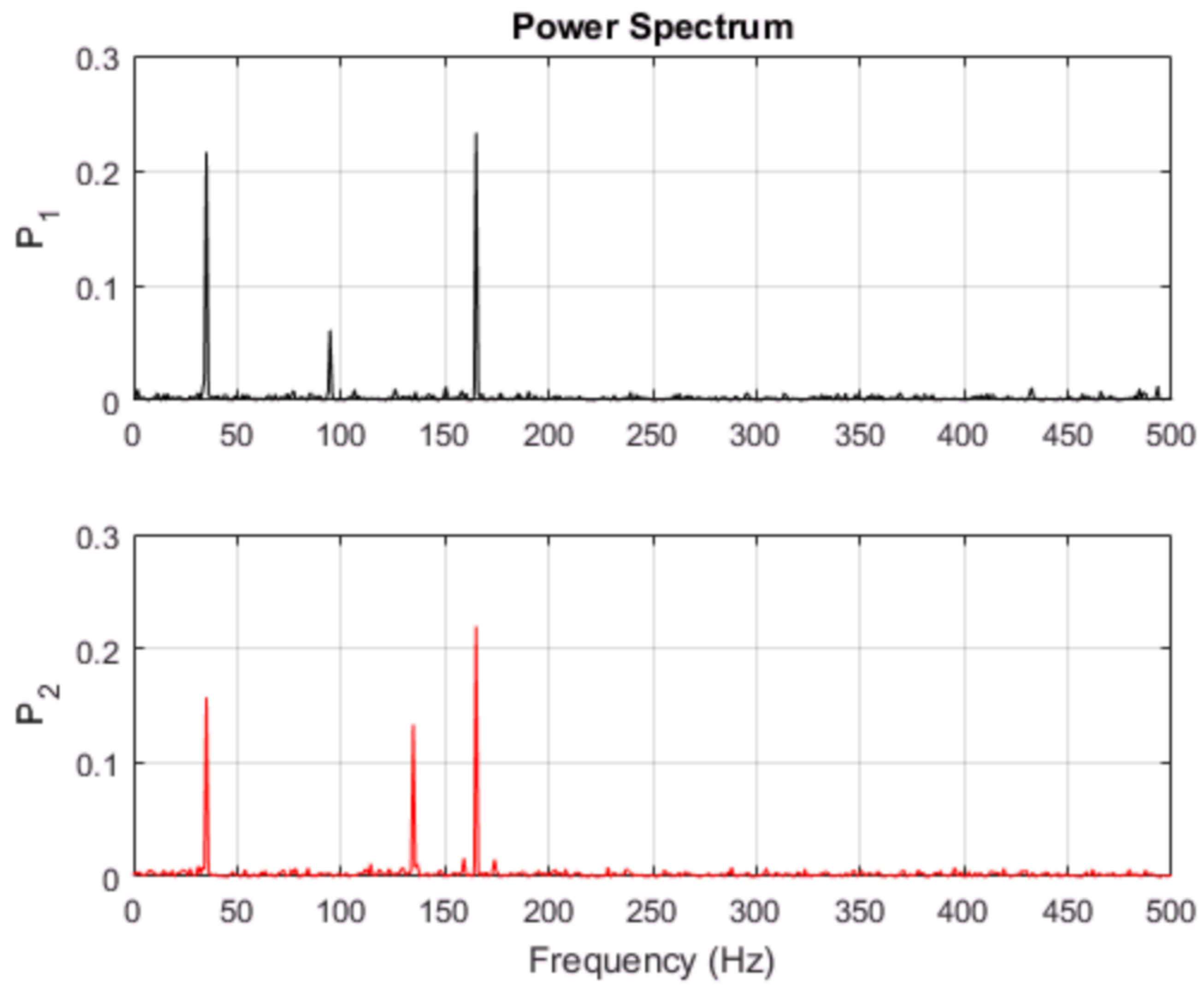
SIGNAL DATA: PREPROCESSING AND EXTRACTING FEATURES

Simple Vibration Analysis in the Time Domain



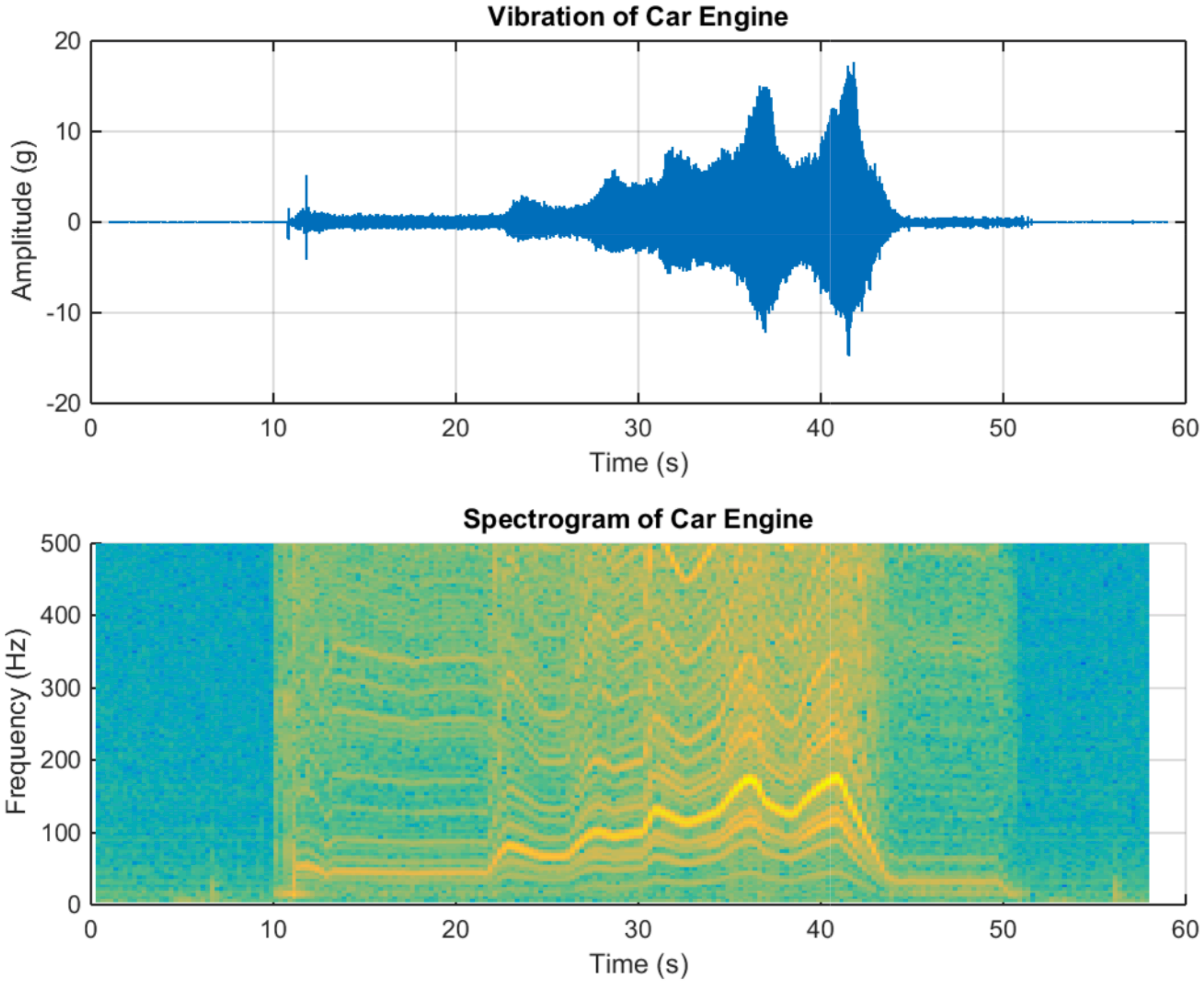
SIGNAL DATA: PREPROCESSING AND EXTRACTING FEATURES

Simple Vibration Analysis in the Frequency Domain



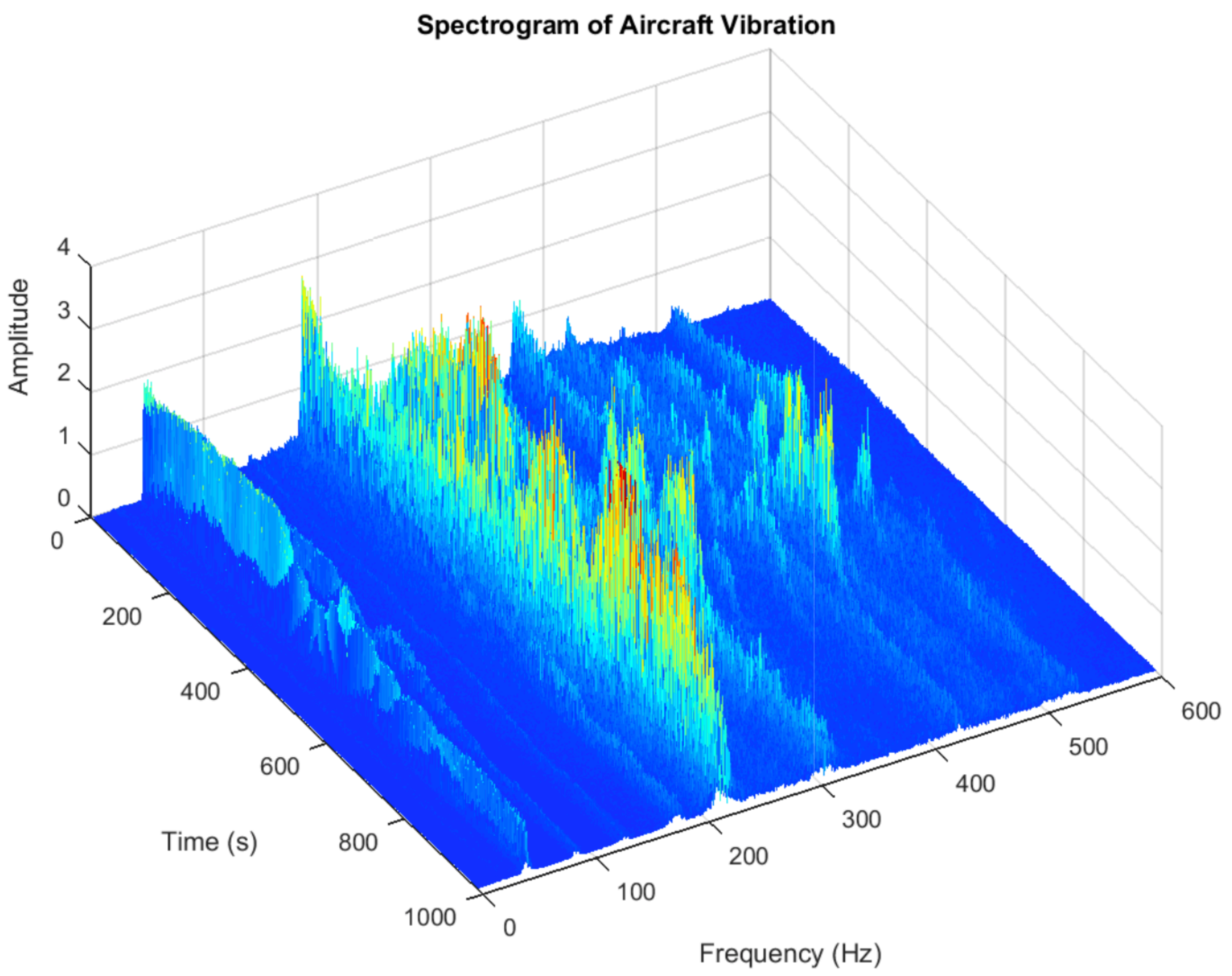
SIGNAL DATA: PREPROCESSING AND EXTRACTING FEATURES

Simple Vibration Analysis in the Time-Frequency Domain



SIGNAL DATA: PREPROCESSING AND EXTRACTING FEATURES

Simple Vibration Analysis in the Time-Frequency Domain



Example: Equipment Fault Diagnostics – Telemetry Data

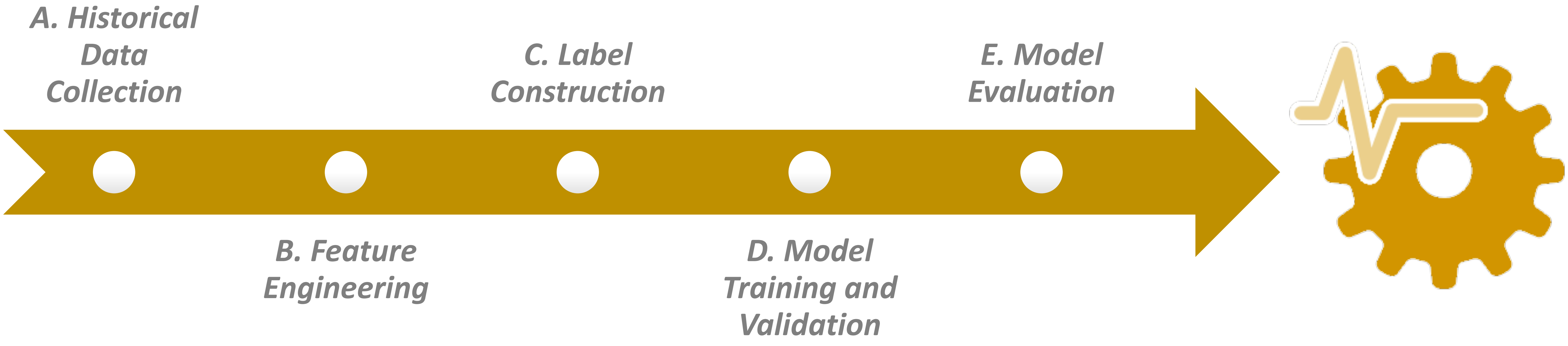
Example: Equipment Fault Diagnostics – Telemetry Data

<u>*Problem Description:*</u>

- The business problem for this example is to answer:

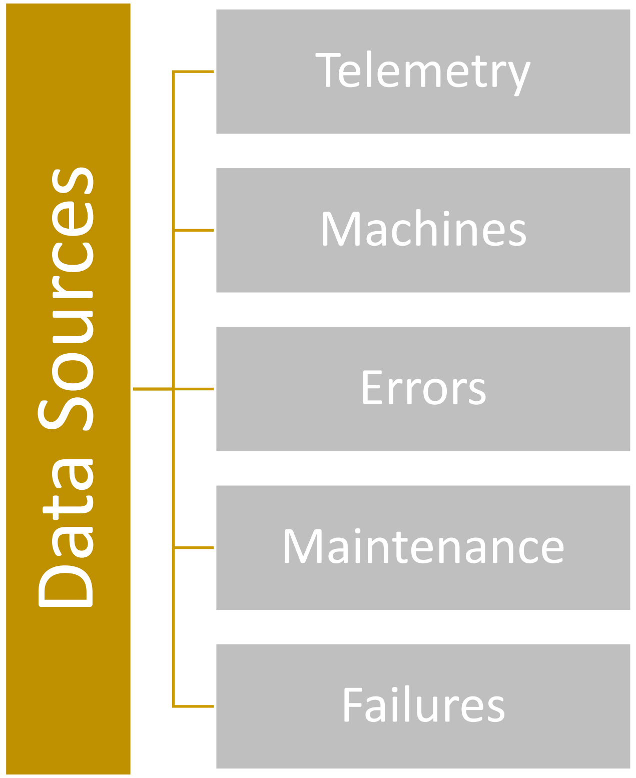
What is the probability that a machine will fail in the near future due to a failure of a certain component?

- In the following sections, we go through the steps of developing a PHM diagnostics model



Example: Equipment Fault Diagnostics – Telemetry Data

<u>A. Historical Data Collection:</u>



Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

Data Sources
- Telemetry
- Machines
- Errors
- Maintenance
- Failures

2015 Hourly Data:
365 days X 24 hours/day
= **8760 hours/machine**

1000 machines

Voltage

Rotation Speed

Pressure

Vibration

Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

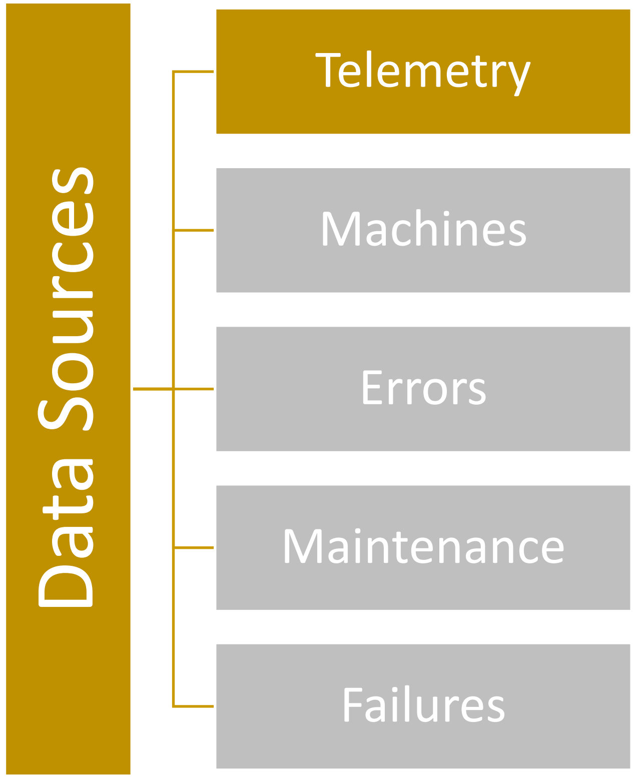
- This data set consists of voltage, rotation, pressure and vibration measurements collected from 1000 machines.
- Total Number of telemetry records: 8761000.

Data Sources
- **Telemetry**
- Machines
- Errors
- Maintenance
- Failures

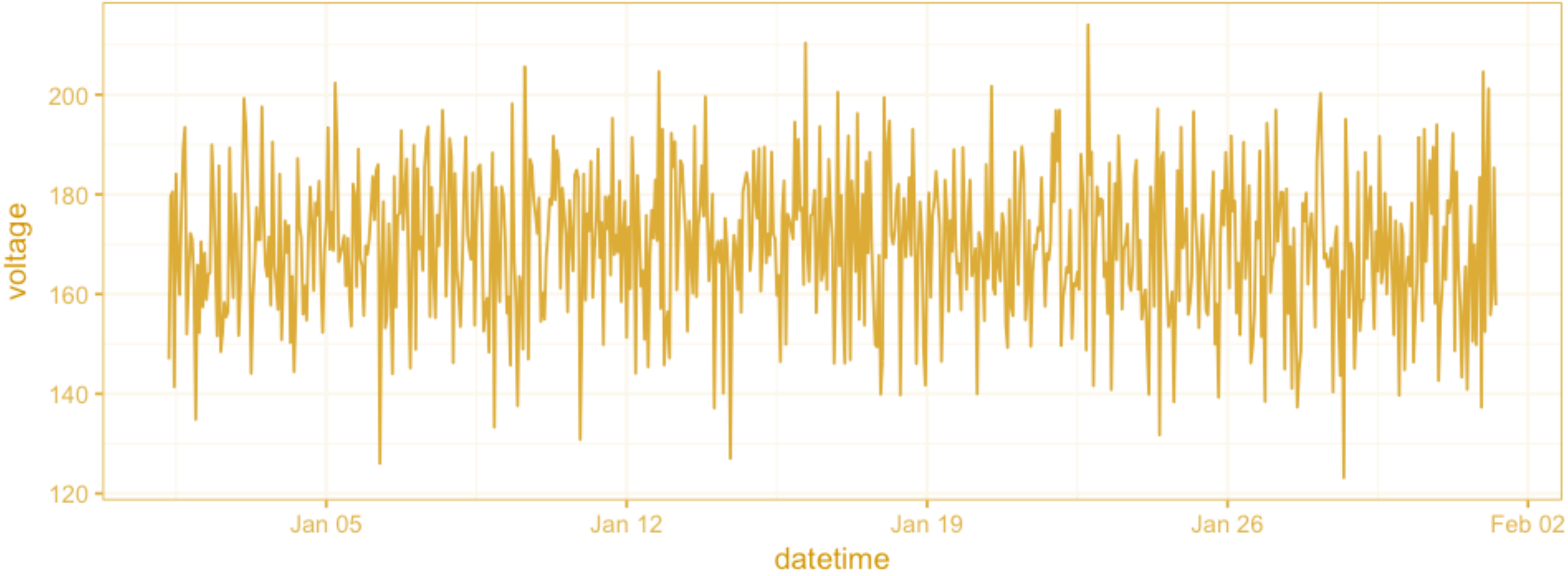
datetime	machineID	volt	rotate	pressure	vibration
1/1/15 6:00	1	151.9200	530.8136	101.78818	49.60401
1/1/15 7:00	1	174.5220	535.5235	113.25601	41.51591
1/1/15 8:00	1	146.9128	456.0807	107.78696	42.09969
1/1/15 9:00	1	179.5306	503.4700	108.28382	37.84773
1/1/15 10:00	1	180.5443	371.6006	107.55331	41.46788
1/1/15 11:00	1	141.4118	530.8573	87.61400	44.98585
1/1/15 12:00	1	184.0838	450.2275	87.69738	30.83126
1/1/15 13:00	1	166.6326	486.4668	108.06773	50.38005
1/1/15 14:00	1	159.8927	488.9687	102.13188	43.66130
1/1/15 15:00	1	176.6868	508.2028	90.95119	43.03970

Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:



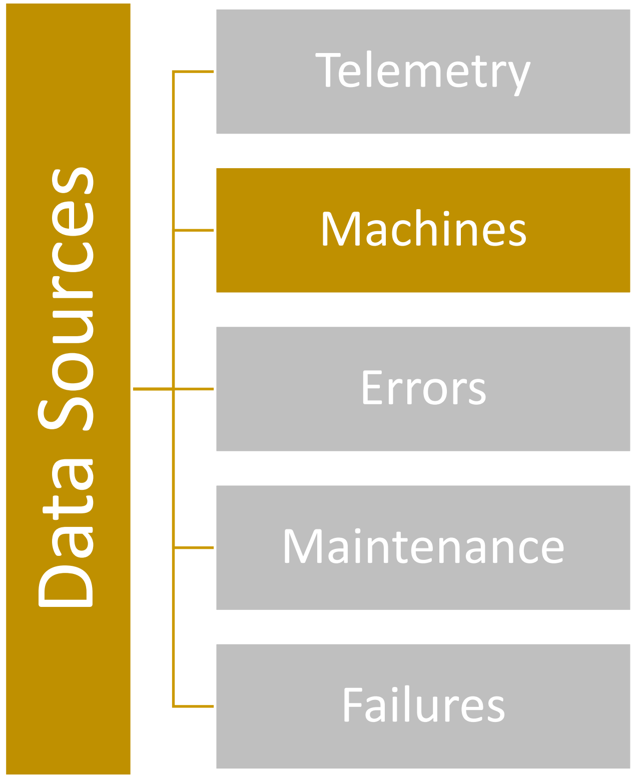
Voltage for MachineID 1 – January 2015



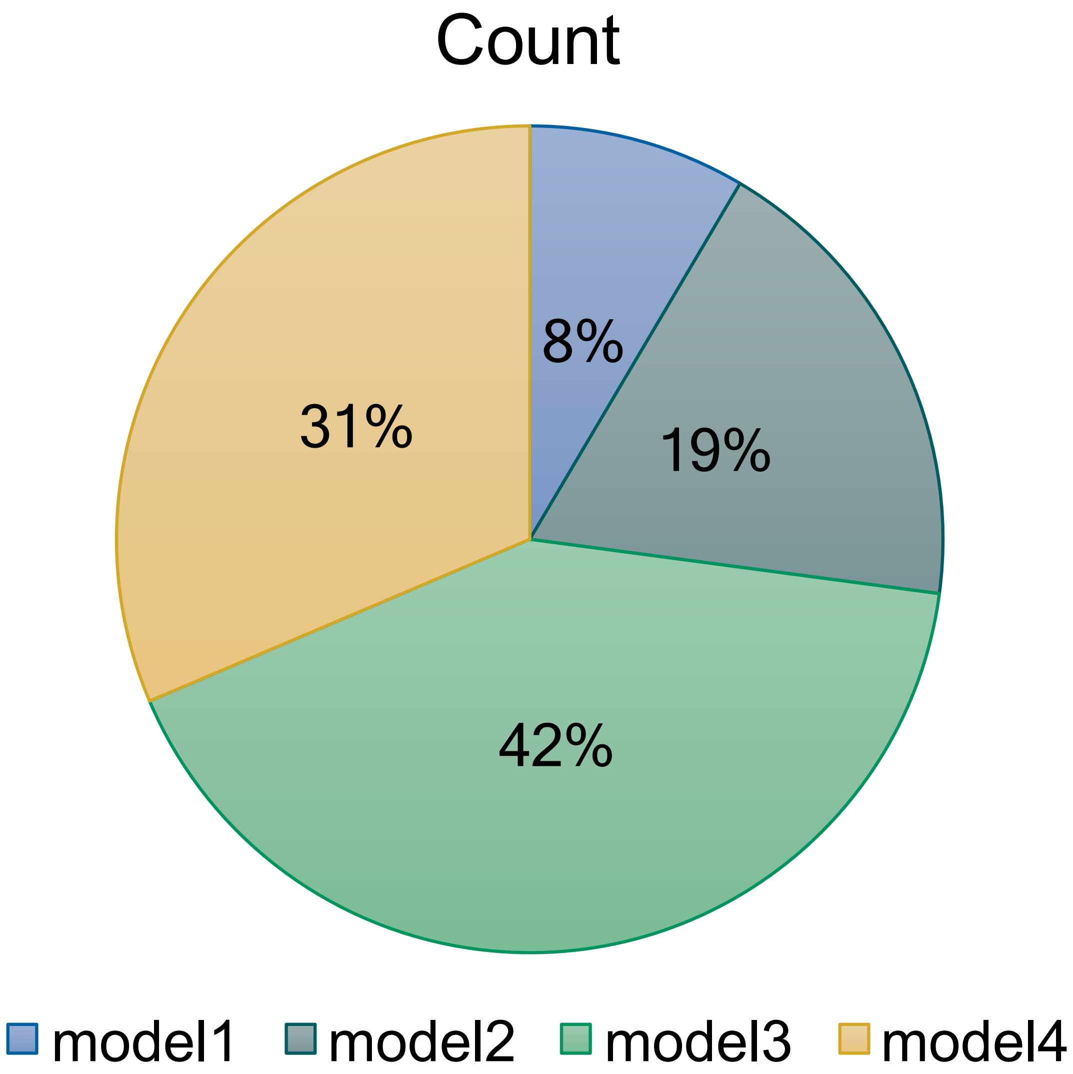
Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

- This data set includes information about the machines which are model type and age which is years in service.
- Total Number of machines: 1000.



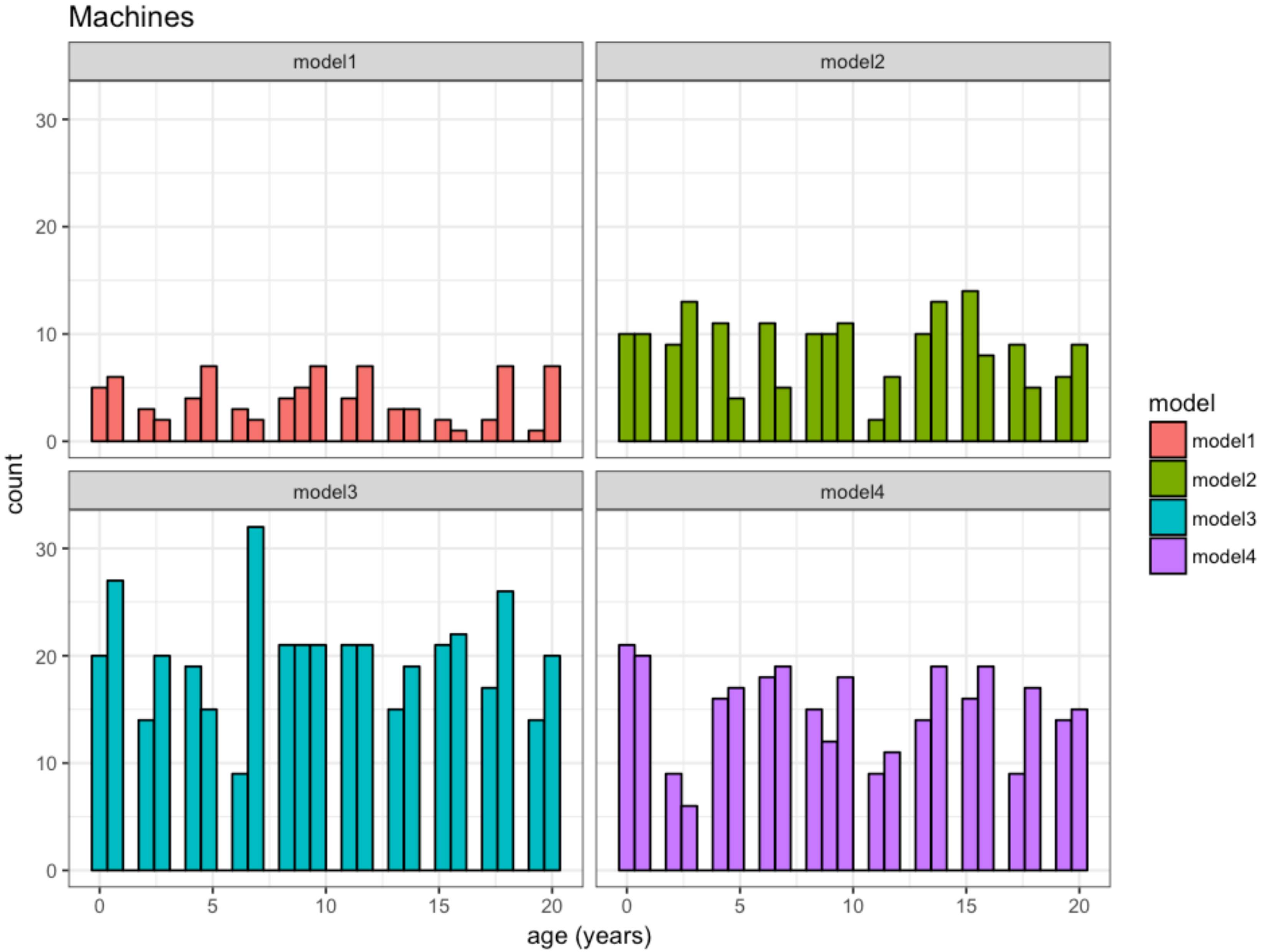
machineID	model	age
1	model2	18
2	model4	7
3	model3	8
4	model3	7
5	model2	2
. . .	. . .	. . .
996	model2	20
997	model2	4
998	model2	14
999	model3	0
1000	model4	16



Example: Equipment Fault Diagnostics – Telemetry Data

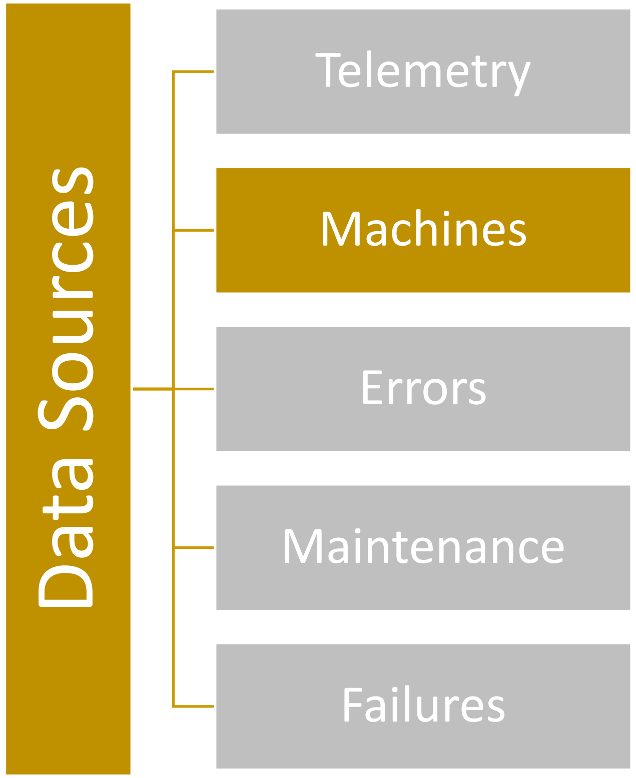
A. Historical Data Collection:

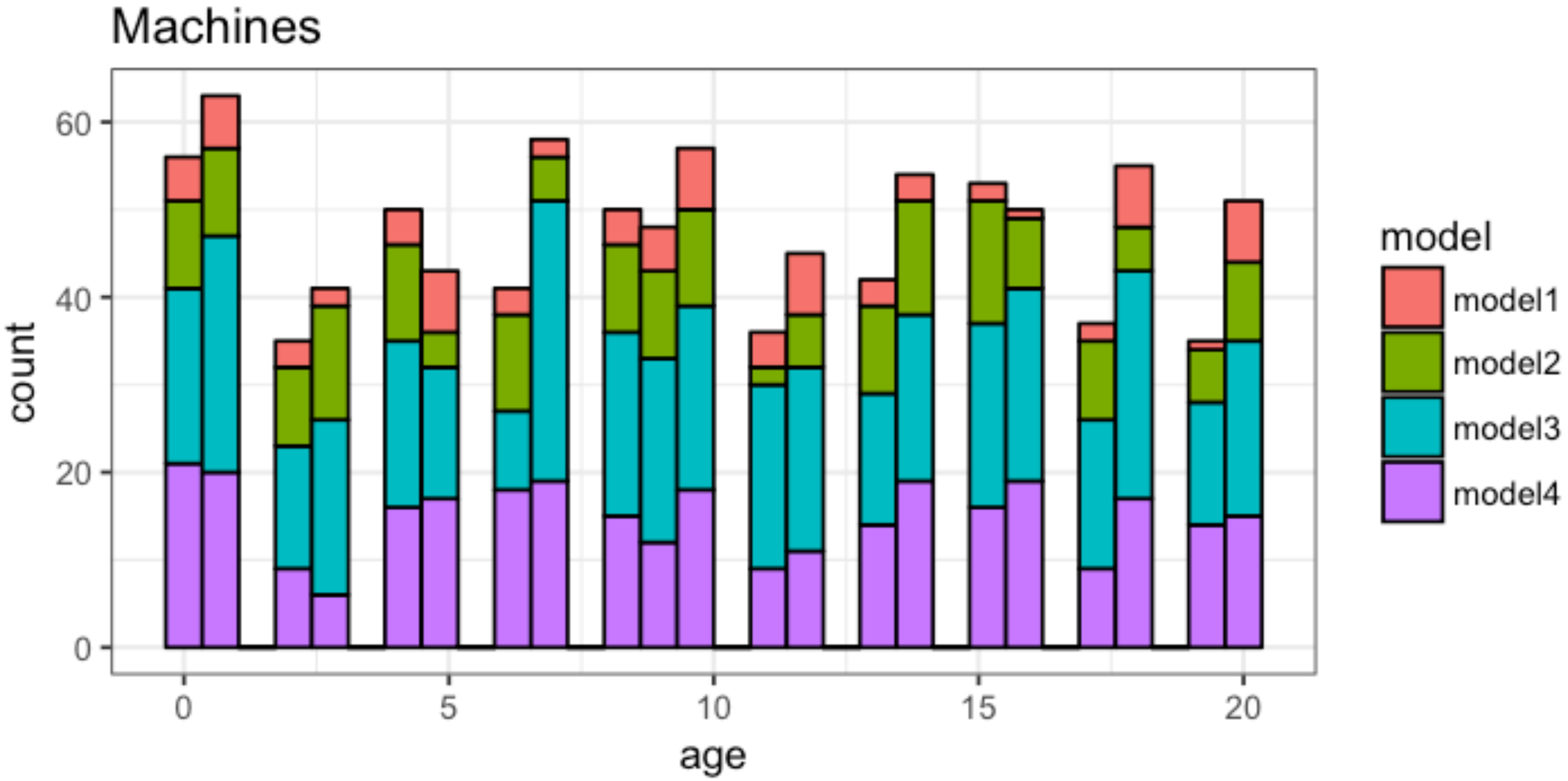
Data Sources
- Telemetry
- **Machines**
- Errors
- Maintenance
- Failures



Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

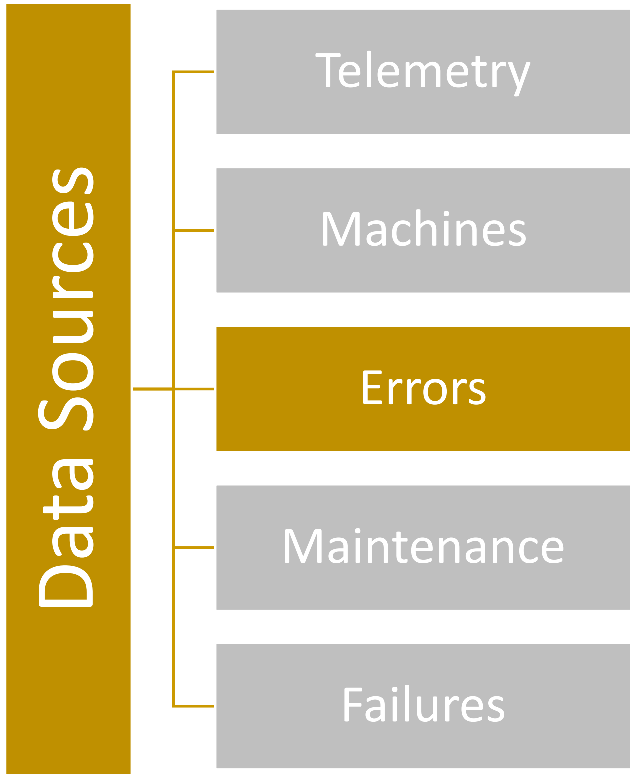




Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

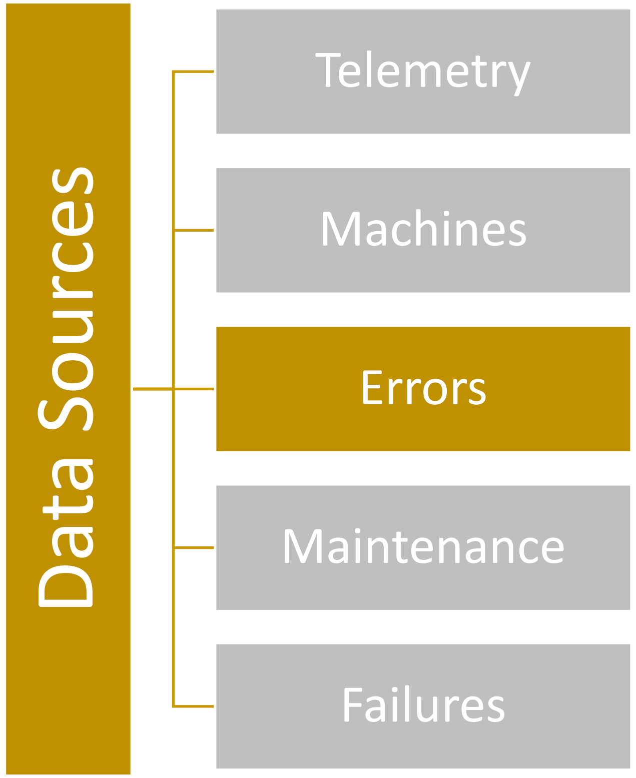
- These are non-breaking errors thrown while the machine is still operational and do not constitute as failures.
- Total Number of error records: 11967.

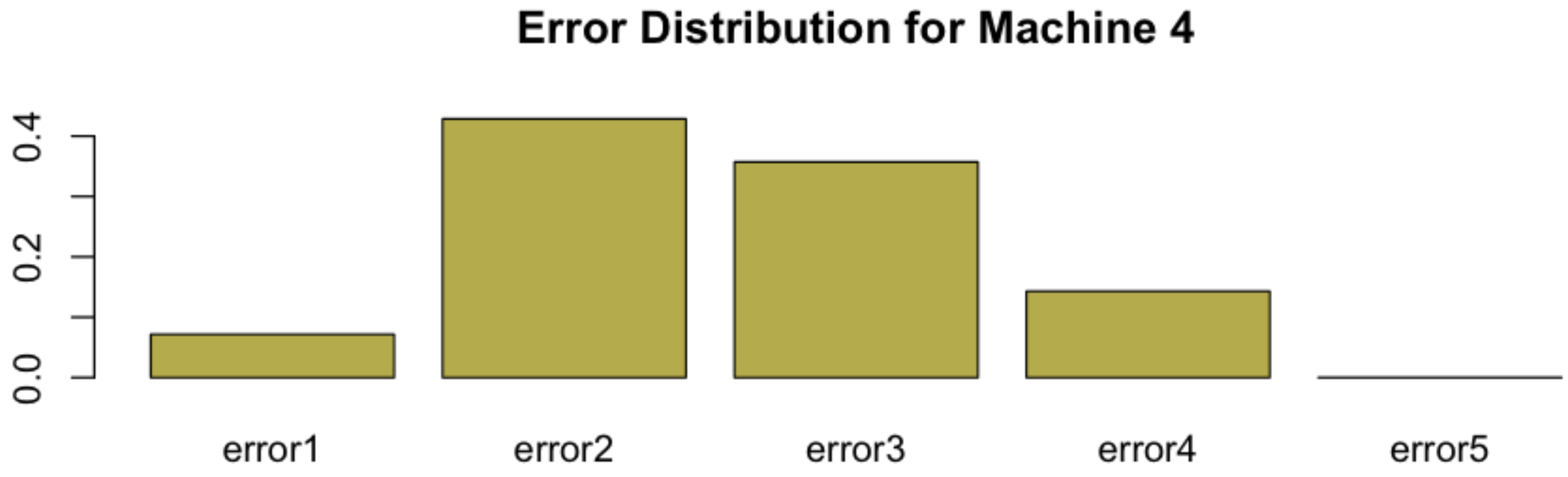


datetime	machineID	errorID
1/6/15 3:00	1	error3
2/3/15 6:00	1	error4
2/21/15 11:00	1	error1
2/21/15 16:00	1	error2
3/20/15 6:00	1	error1
. . .	. . .	. . .
6/20/15 8:00	1000	error2
7/5/15 4:00	1000	error1
8/12/15 6:00	1000	error2
8/12/15 6:00	1000	error3
9/11/15 6:00	1000	error1

Example: Equipment Fault Diagnostics – Telemetry Data

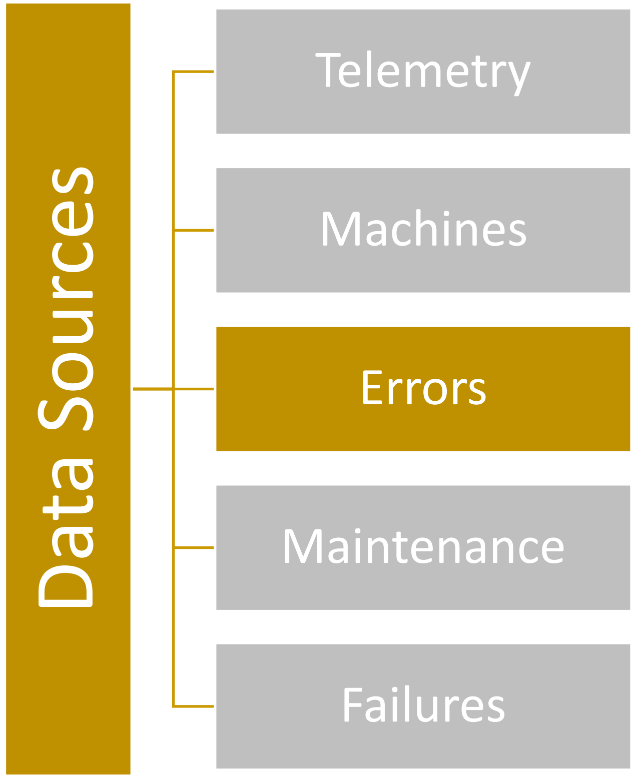
A. Historical Data Collection:

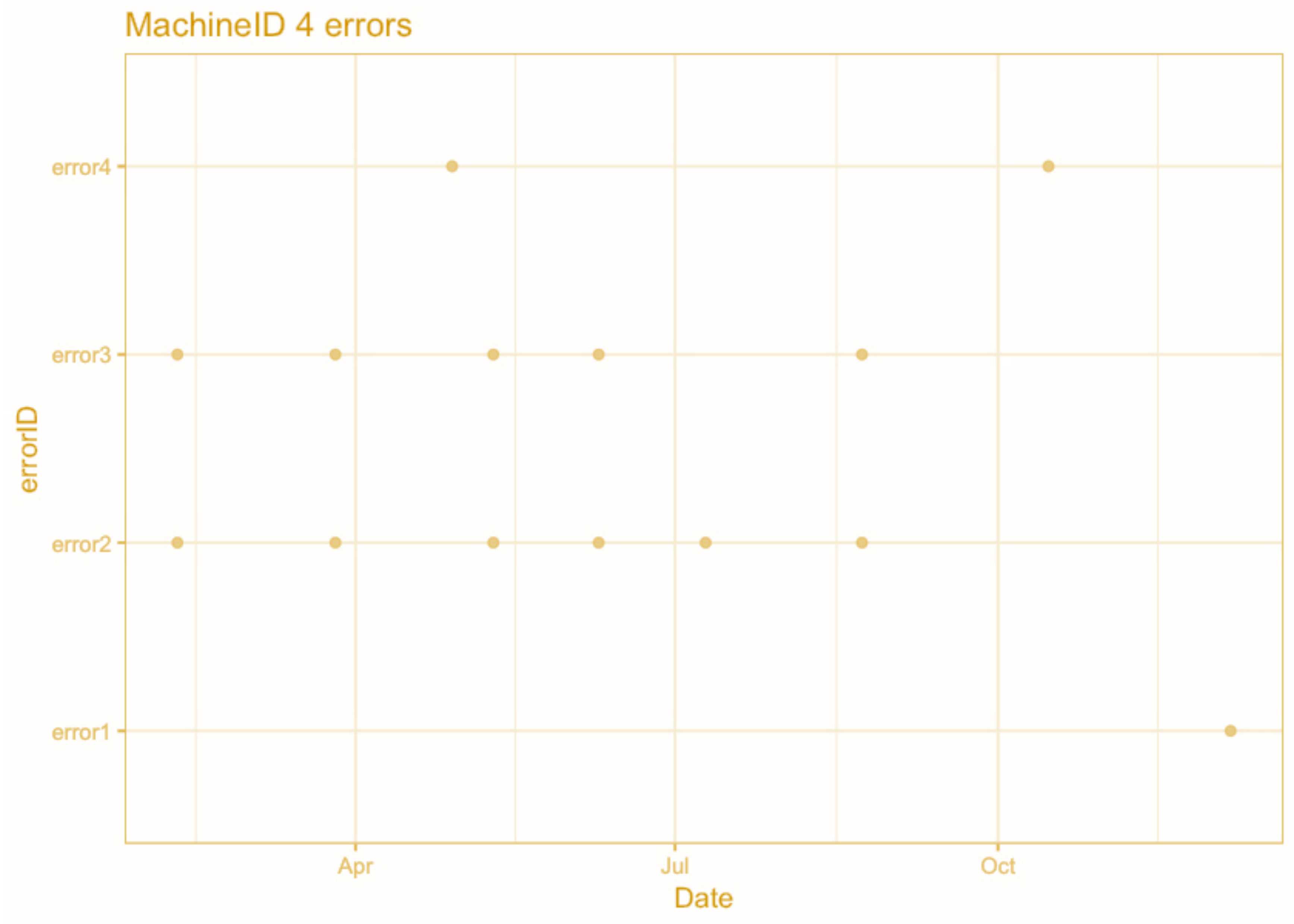




Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

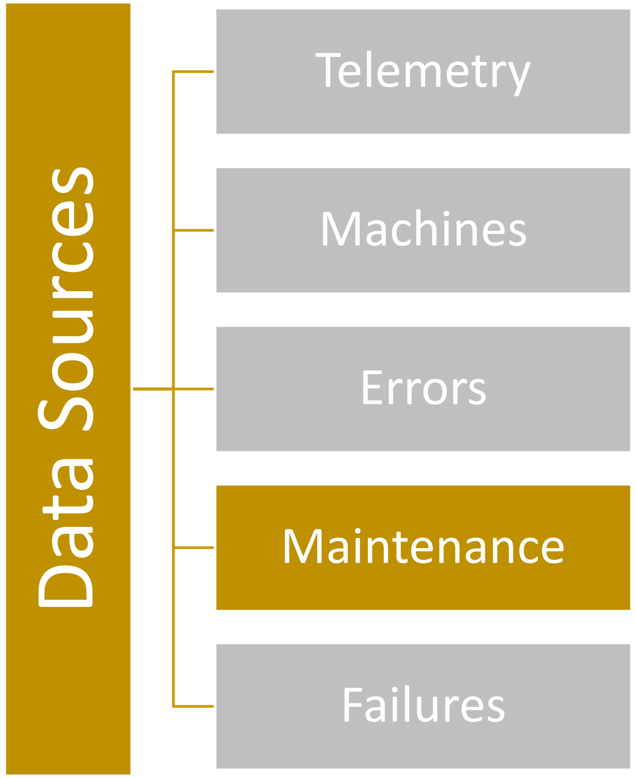




Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

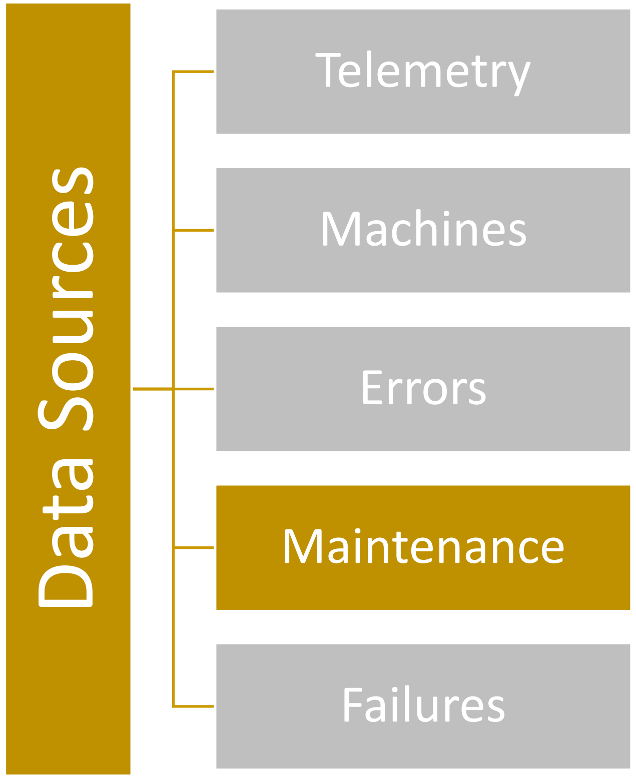
- These are the scheduled and unscheduled maintenance records which correspond to both regular inspection of components as well as failures.
- Total Number of maintenance records: 32592.

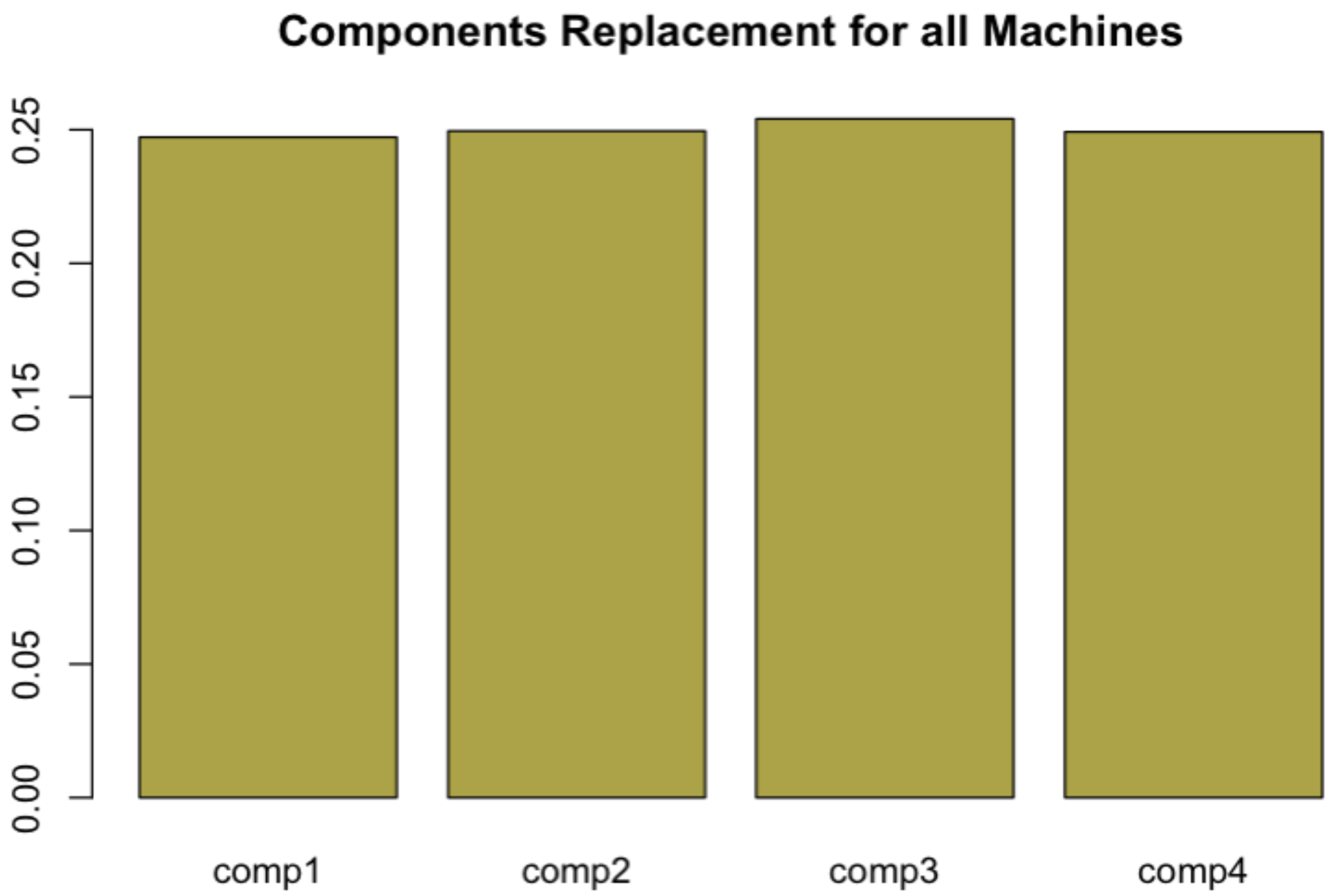


datetime	machineID	comp
7/1/14 6:00	1	comp4
9/14/14 6:00	1	comp1
9/14/14 6:00	1	comp2
11/13/14 6:00	1	comp3
1/5/15 6:00	1	comp1
. . .	. . .	. . .
10/27/15 6:00	1000	comp1
11/11/15 6:00	1000	comp1
12/11/15 6:00	1000	comp1
12/26/15 6:00	1000	comp3
12/26/15 6:00	1000	comp1

Example: Equipment Fault Diagnostics – Telemetry Data

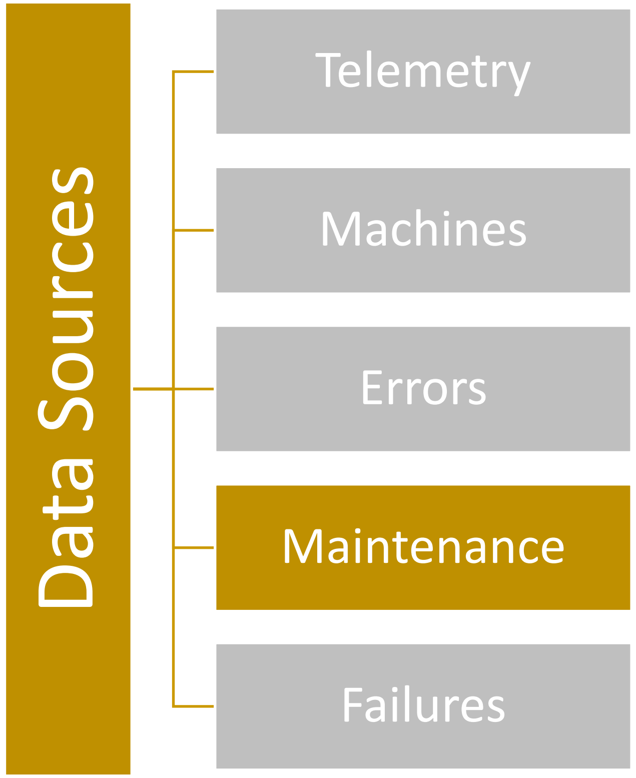
A. Historical Data Collection:

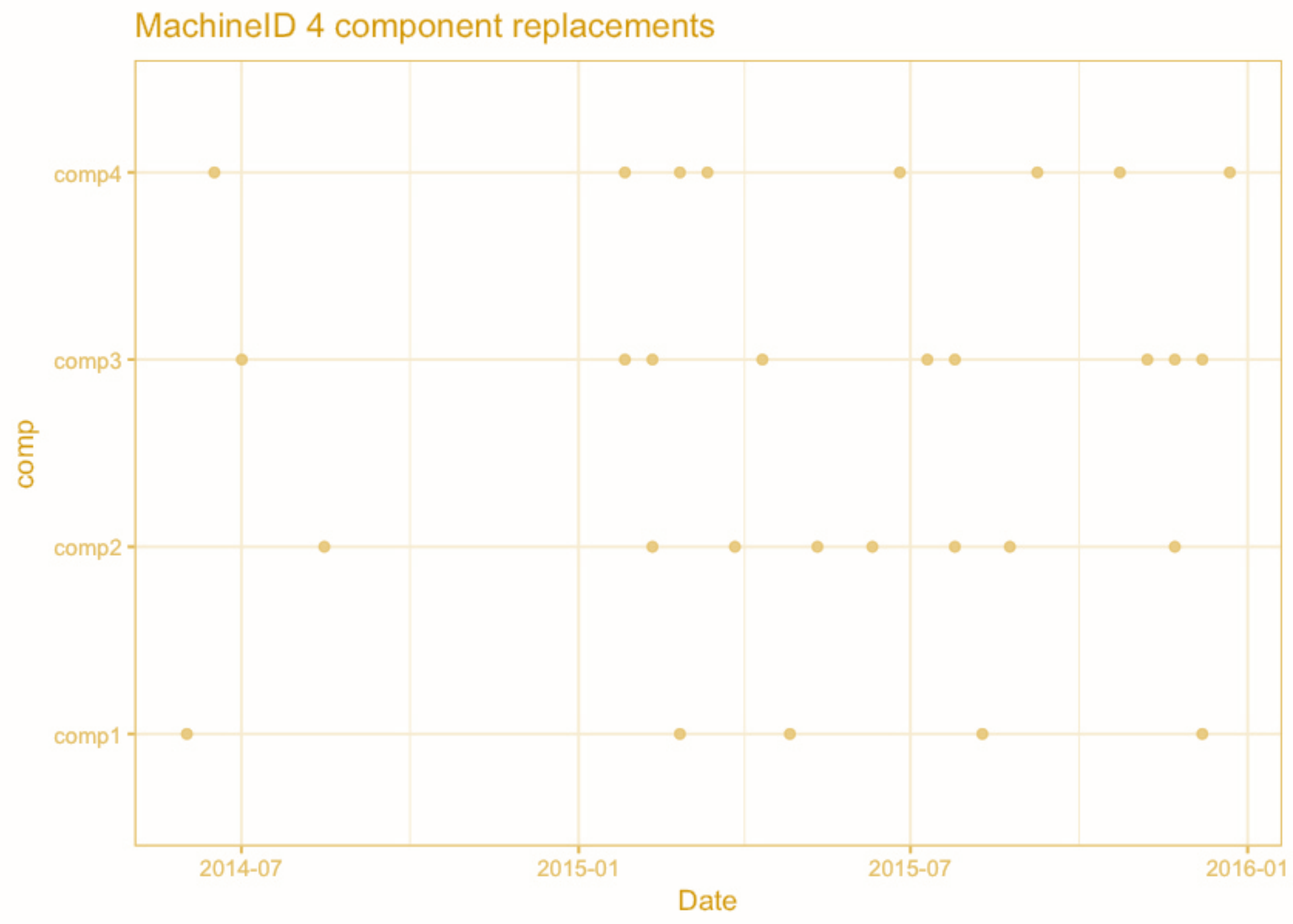




Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

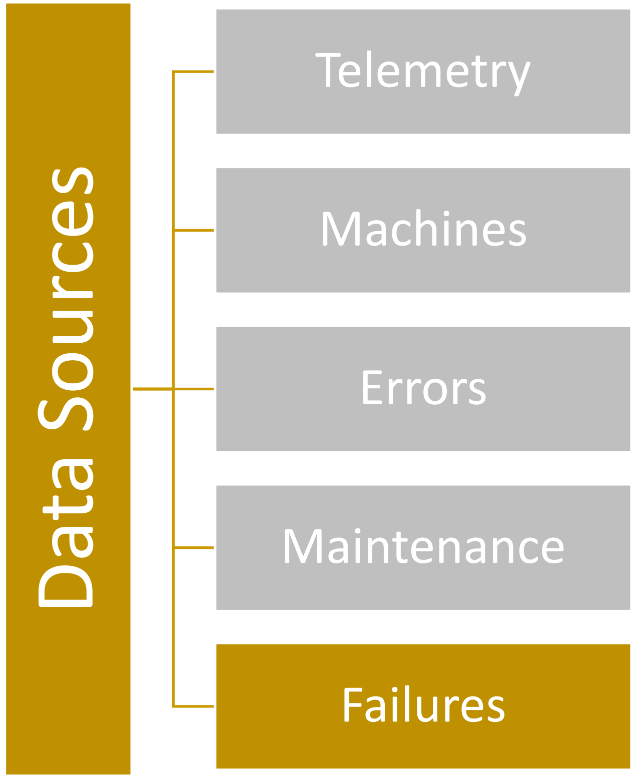




Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:

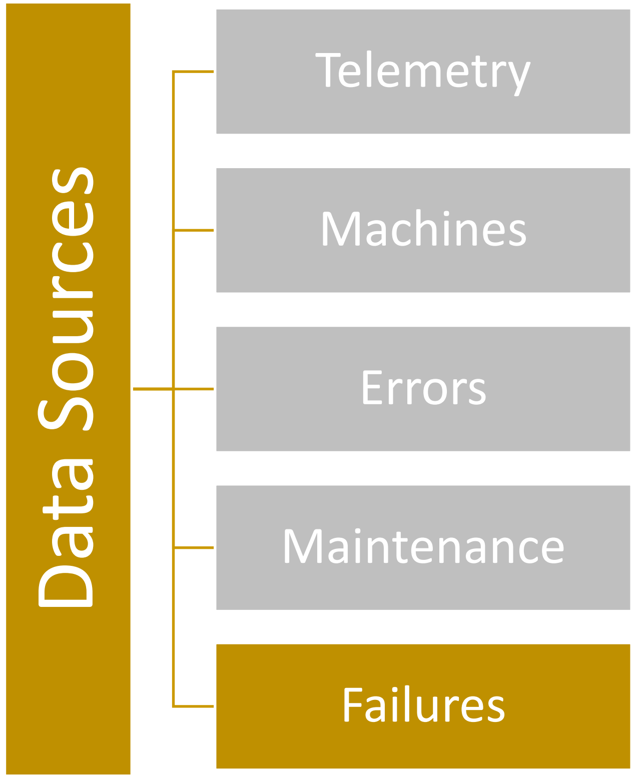
- These are the records of component replacements due to failures.
- Total Number of failures: 6726.

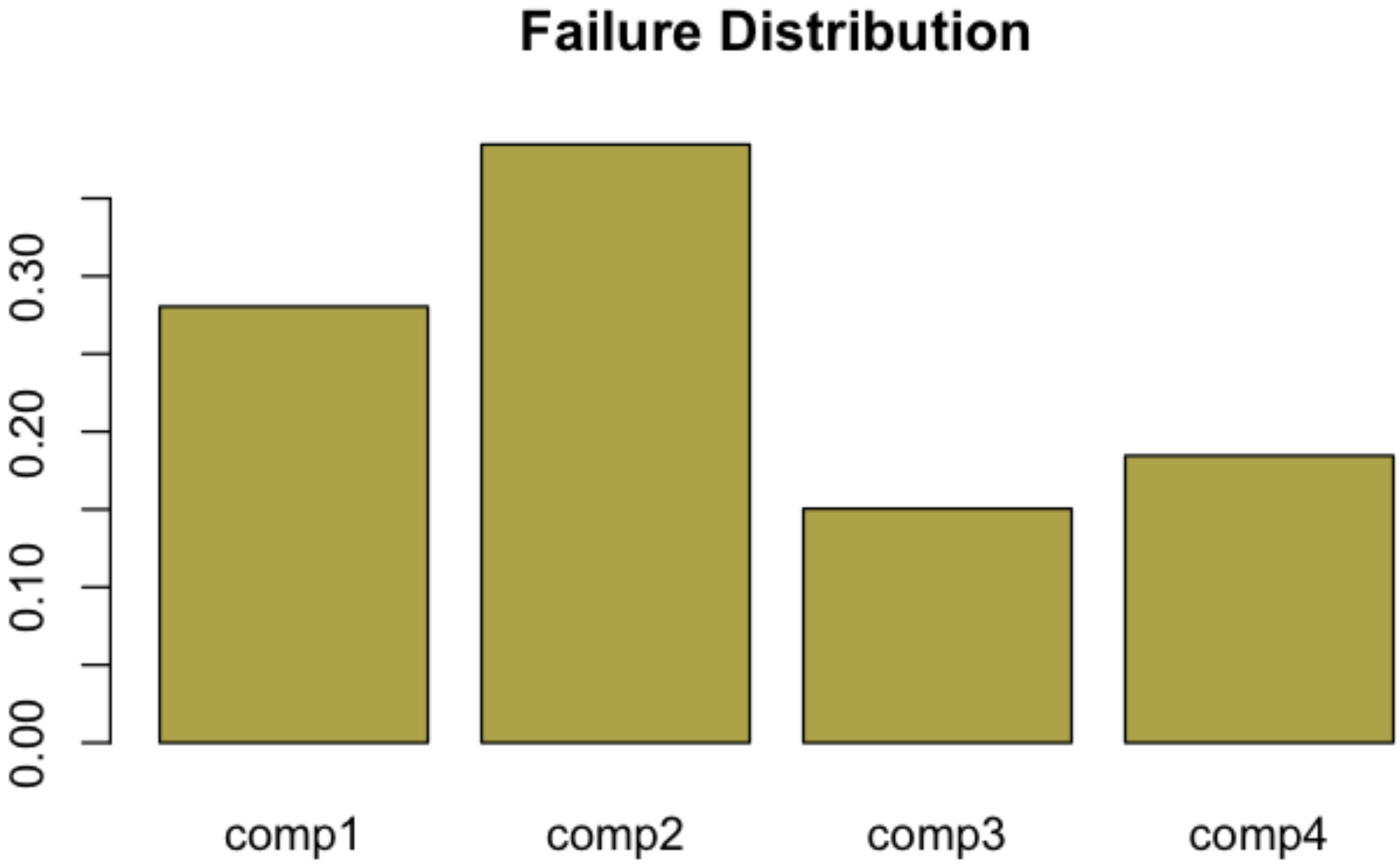


datetime	machineID	failure
2/4/15 6:00	1	comp3
3/21/15 6:00	1	comp1
4/5/15 6:00	1	comp4
5/5/15 6:00	1	comp3
5/20/15 6:00	1	comp2
⋮	⋮	⋮
1/30/15 6:00	1000	comp4
3/31/15 6:00	1000	comp4
5/15/15 6:00	1000	comp2
8/13/15 6:00	1000	comp2
9/12/15 6:00	1000	comp1

Example: Equipment Fault Diagnostics – Telemetry Data

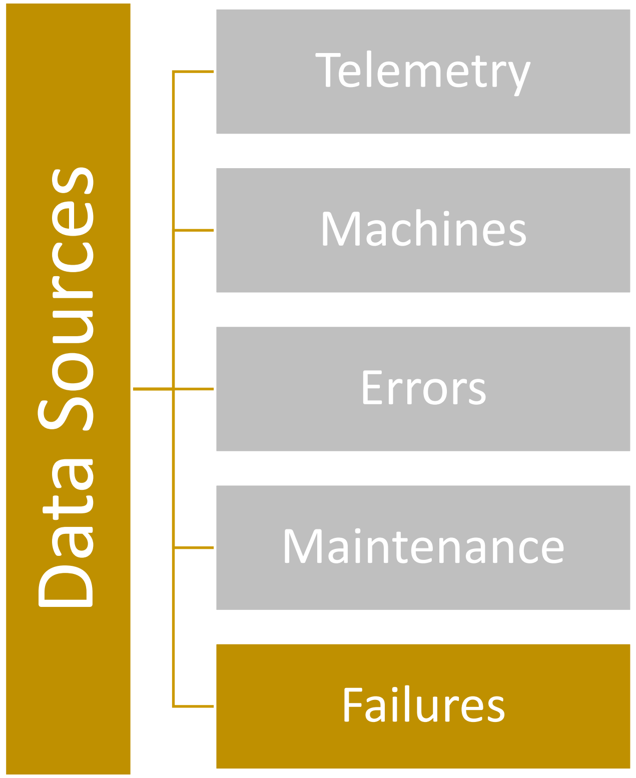
A. Historical Data Collection:

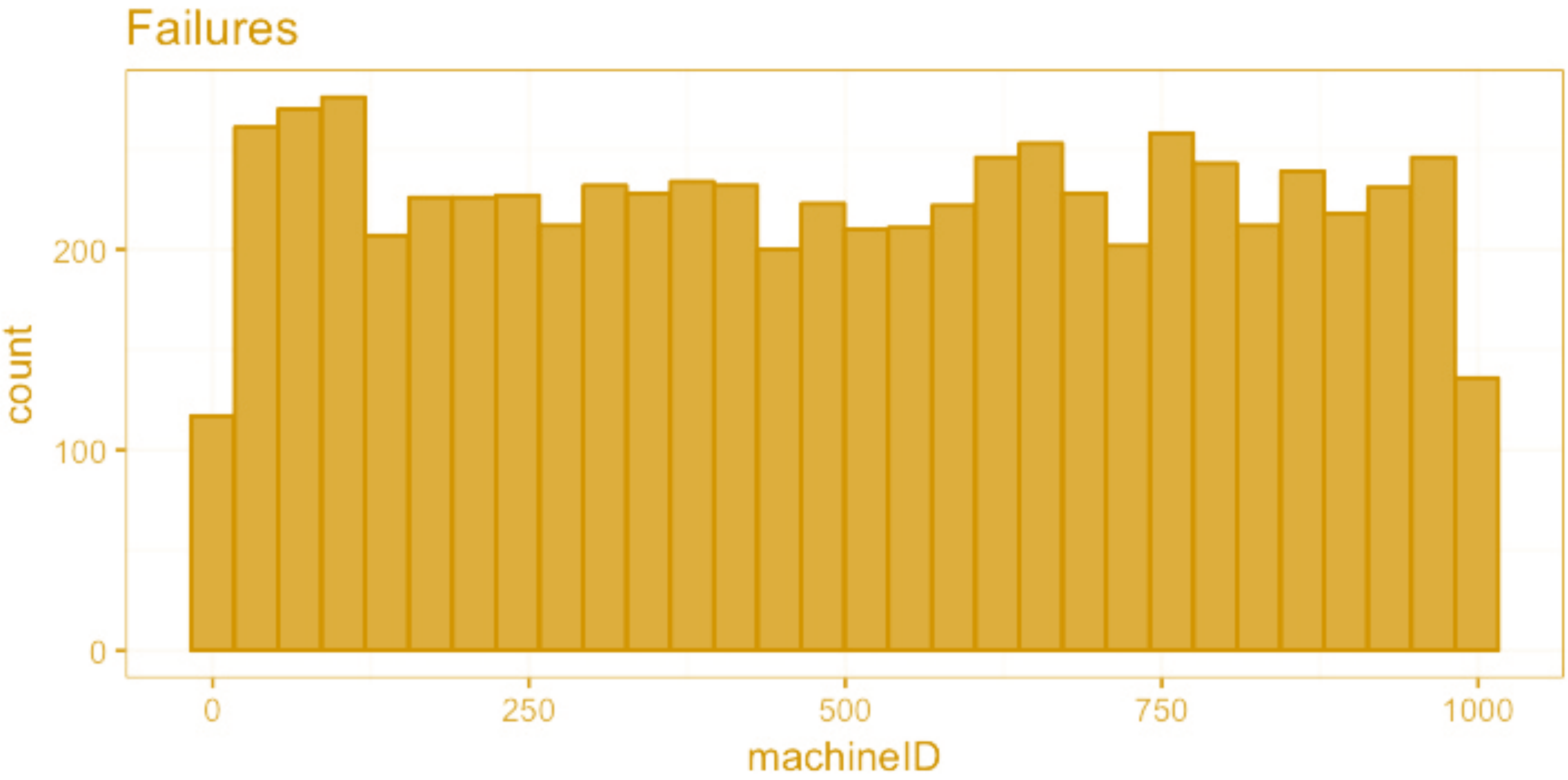




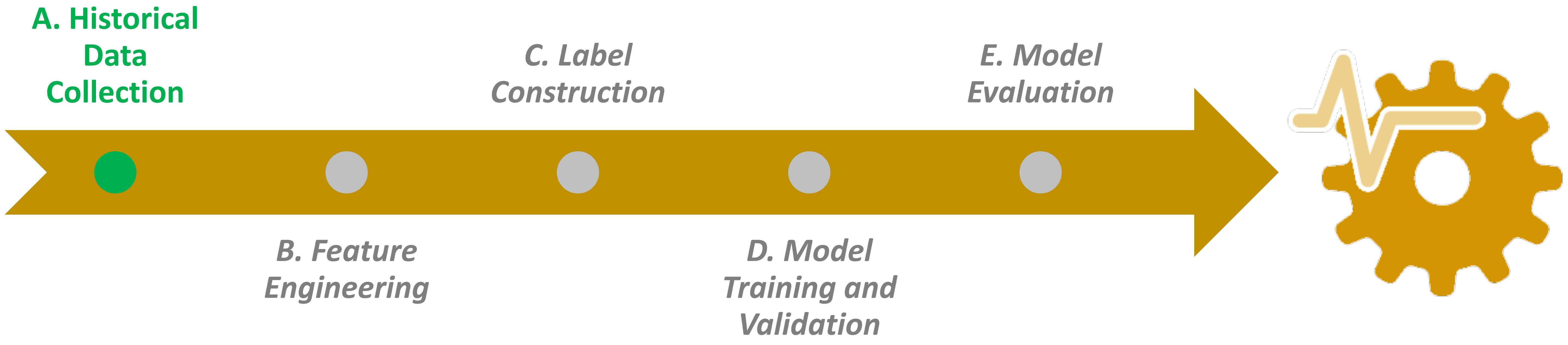
Example: Equipment Fault Diagnostics – Telemetry Data

A. Historical Data Collection:



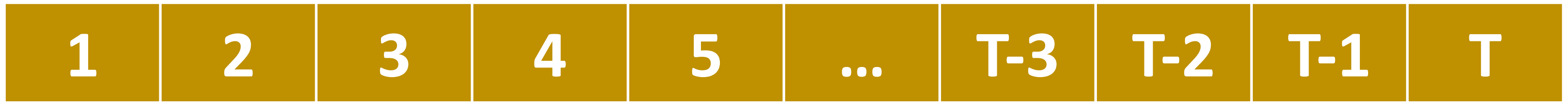


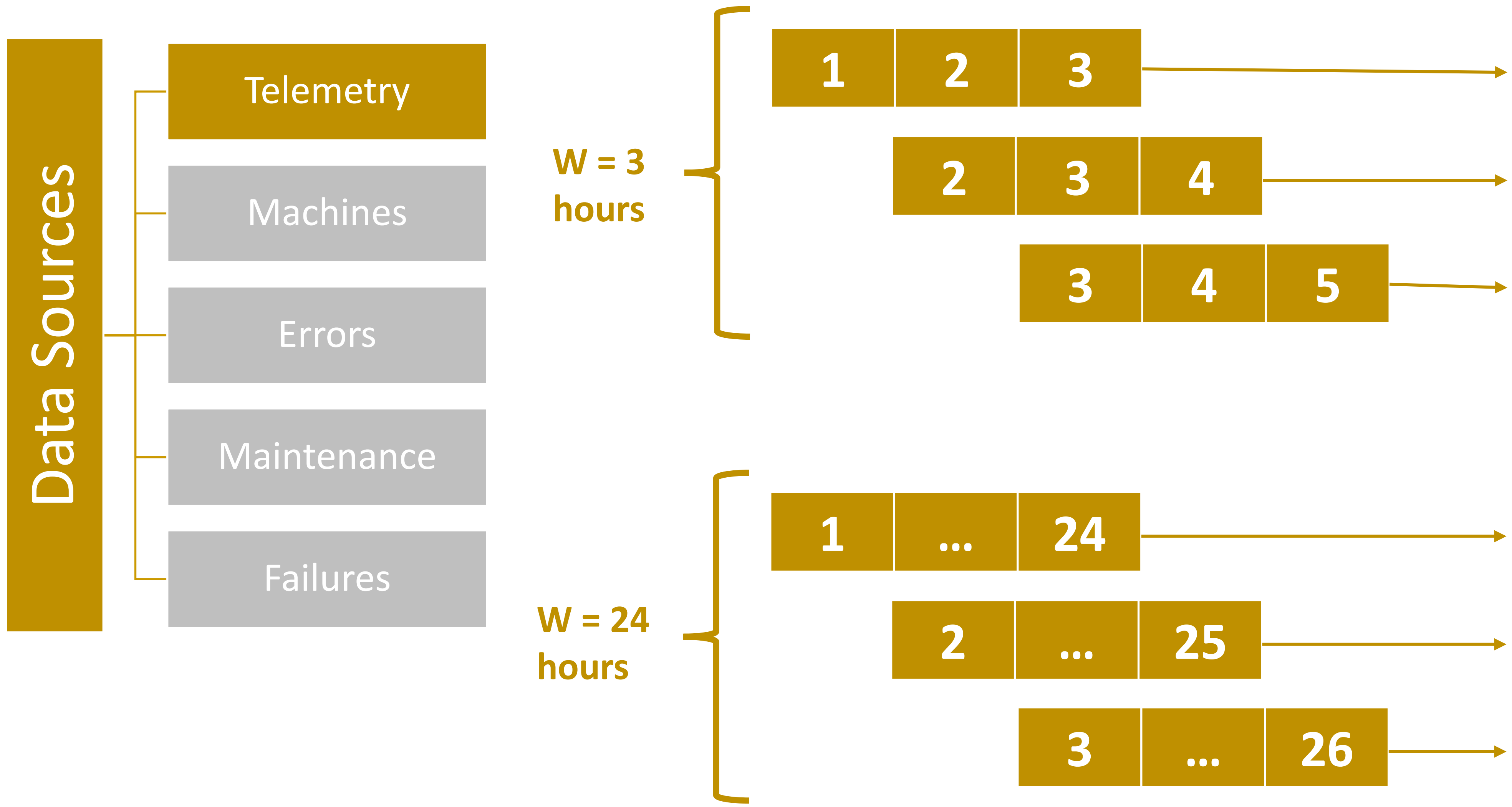
Example: Equipment Fault Diagnostics – Telemetry Data



Example: Equipment Fault Diagnostics – Telemetry Data

B. Feature Engineering:



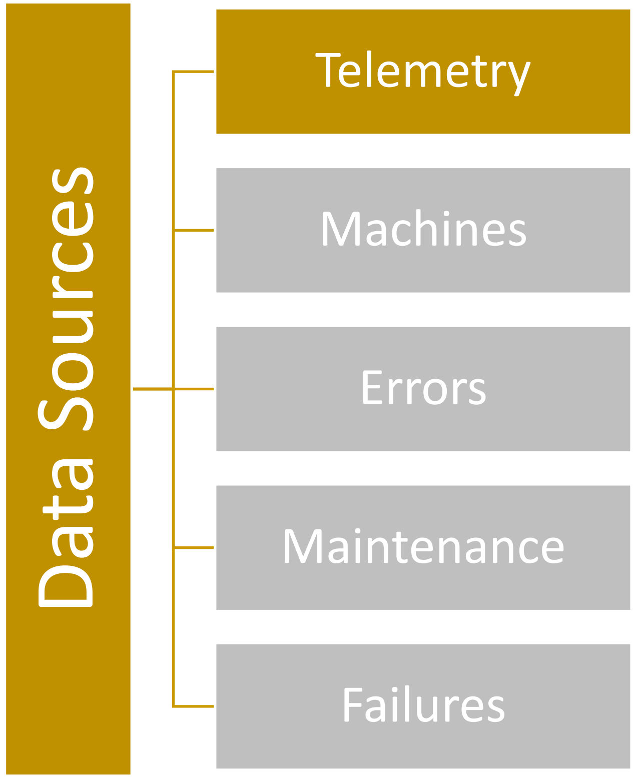


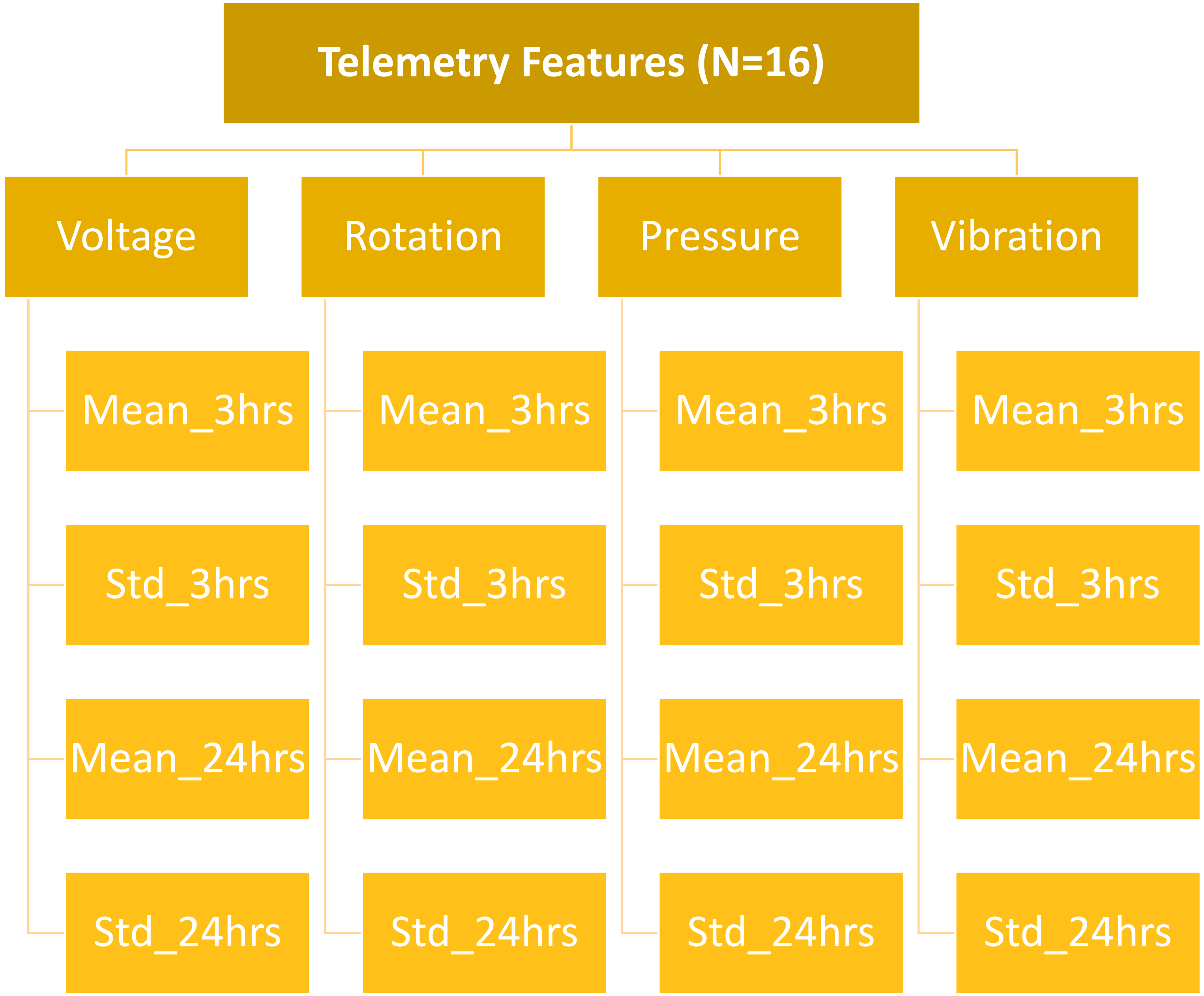
Mean	Standard Deviation
u1	s1
u2	s2
u3	s3

Mean_24hrs	Standard Deviation_24hrs
u1	s1
u2	s2
u3	s3

Example: Equipment Fault Diagnostics – Telemetry Data

B. Feature Engineering:





Example: Equipment Fault Diagnostics – Telemetry Data

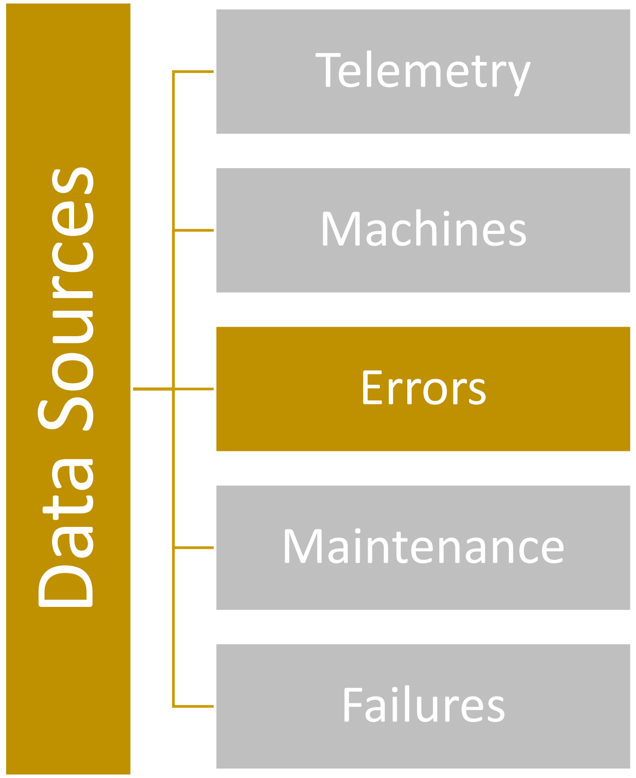
B. Feature Engineering:

- Data Sources
  - Telemetry
  - Machines
  - Errors
  - Maintenance
  - Failures

- **Machine Features (N=2)**
  - Model
  - Age

Example: Equipment Fault Diagnostics – Telemetry Data

B. Feature Engineering:



datetime	errorID
1/6/15 3:00	error3
2/3/15 6:00	error4
2/21/15 11:00	error1
2/21/15 16:00	error2
3/20/15 6:00	error1
. . .	. . .
6/20/15 8:00	error2
7/5/15 4:00	error1
8/12/15 6:00	error2
8/12/15 6:00	error3
9/11/15 6:00	error1

Aggregate multiple errors within an hour

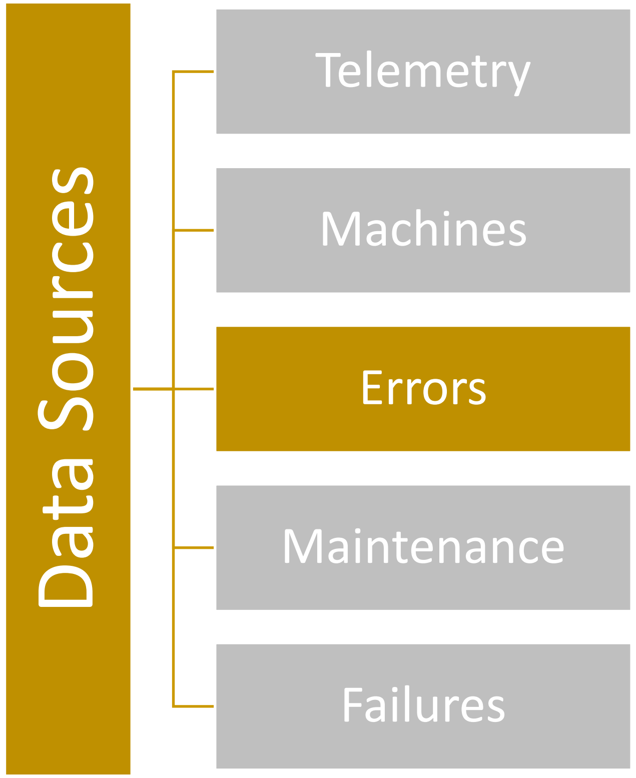
error1	error2	error3	error4	error5
0	0	1	0	0
0	0	0	1	0
1	0	0	0	0
0	1	0	0	0
1	0	0	0	0
. . .	. . .	. . .	. . .	. . .
0	1	0	0	0
1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
1	0	0	0	0

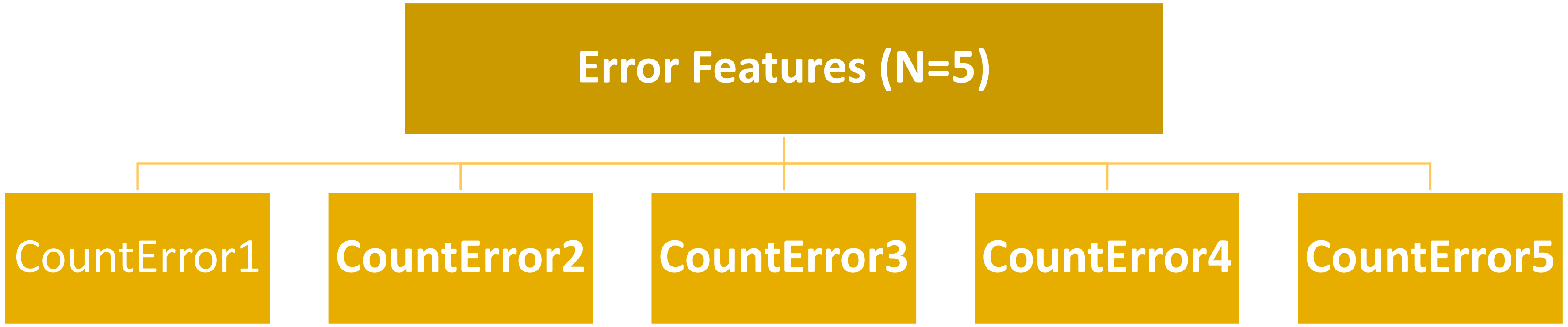
Count the number of errors in the last 24 hours for every 3 hours

Counte rror1	Counter ror2	Counter ror3	Counter ror4	Counter ror5
0	0	1	0	0
0	0	0	2	0
2	0	0	0	0
0	3	0	0	0
1	0	0	0	0
. . .	. . .	. . .	. . .	. . .
0	1	0	0	0
1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
1	0	0	0	0

Example: Equipment Fault Diagnostics – Telemetry Data

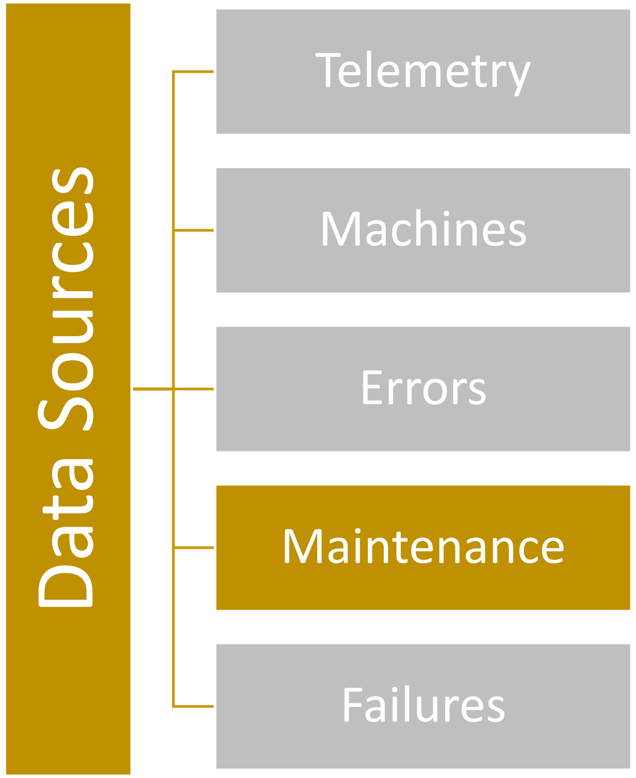
B. Feature Engineering:





Example: Equipment Fault Diagnostics – Telemetry Data

B. Feature Engineering:



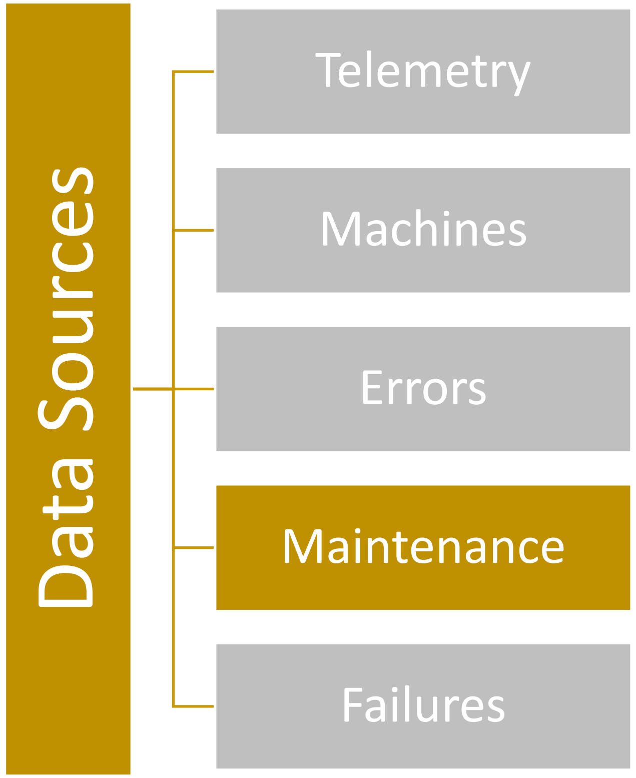
datetime	machineID	comp
7/1/14 6:00	1	comp4
9/14/14 6:00	1	comp1
9/14/14 6:00	1	comp2
11/13/14 6:00	1	comp3
1/5/15 6:00	1	comp1
.	.	.
10/27/15 6:00	1000	comp1
11/11/15 6:00	1000	comp1
12/11/15 6:00	1000	comp1
12/26/15 6:00	1000	comp3
12/26/15 6:00	1000	comp1

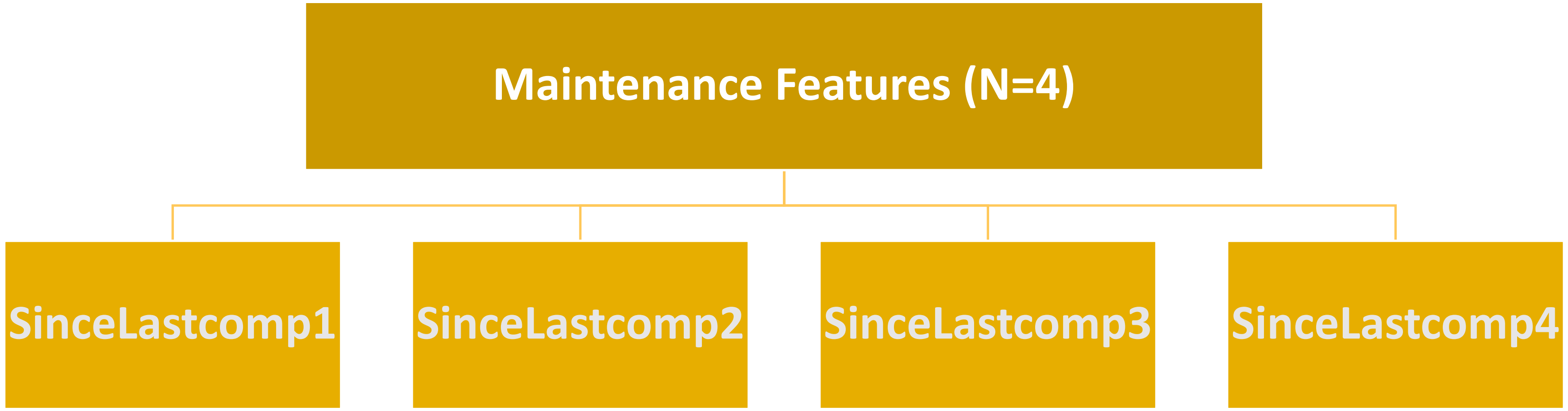
Calculate the time since last replacement

datetime	Machine ID	Since Lastcomp1	Since Lastcomp2	Since Lastcomp3	Since Lastcomp4
1/2/15 5:00	1	109.9583	109.9583	49.95833	184.9583
1/2/15 8:00	1	110.0833	110.0833	50.08333	185.0833
1/2/15 11:00	1	110.2083	110.2083	50.20833	185.2083
1/2/15 14:00	1	110.3333	110.3333	50.33333	185.3333
1/2/15 17:00	1	110.4583	110.4583	50.45833	185.4583
1/2/15 20:00	1	110.5833	110.5833	50.58333	185.5833
1/2/15 23:00	1	110.7083	110.7083	50.70833	185.7083
1/3/15 2:00	1	110.8333	110.8333	50.83333	185.8333
1/3/15 5:00	1	110.9583	110.9583	50.95833	185.9583
1/3/15 8:00	1	111.0833	111.0833	51.08333	186.0833
1/2/15 5:00	1	109.9583	109.9583	49.95833	184.9583

Example: Equipment Fault Diagnostics – Telemetry Data

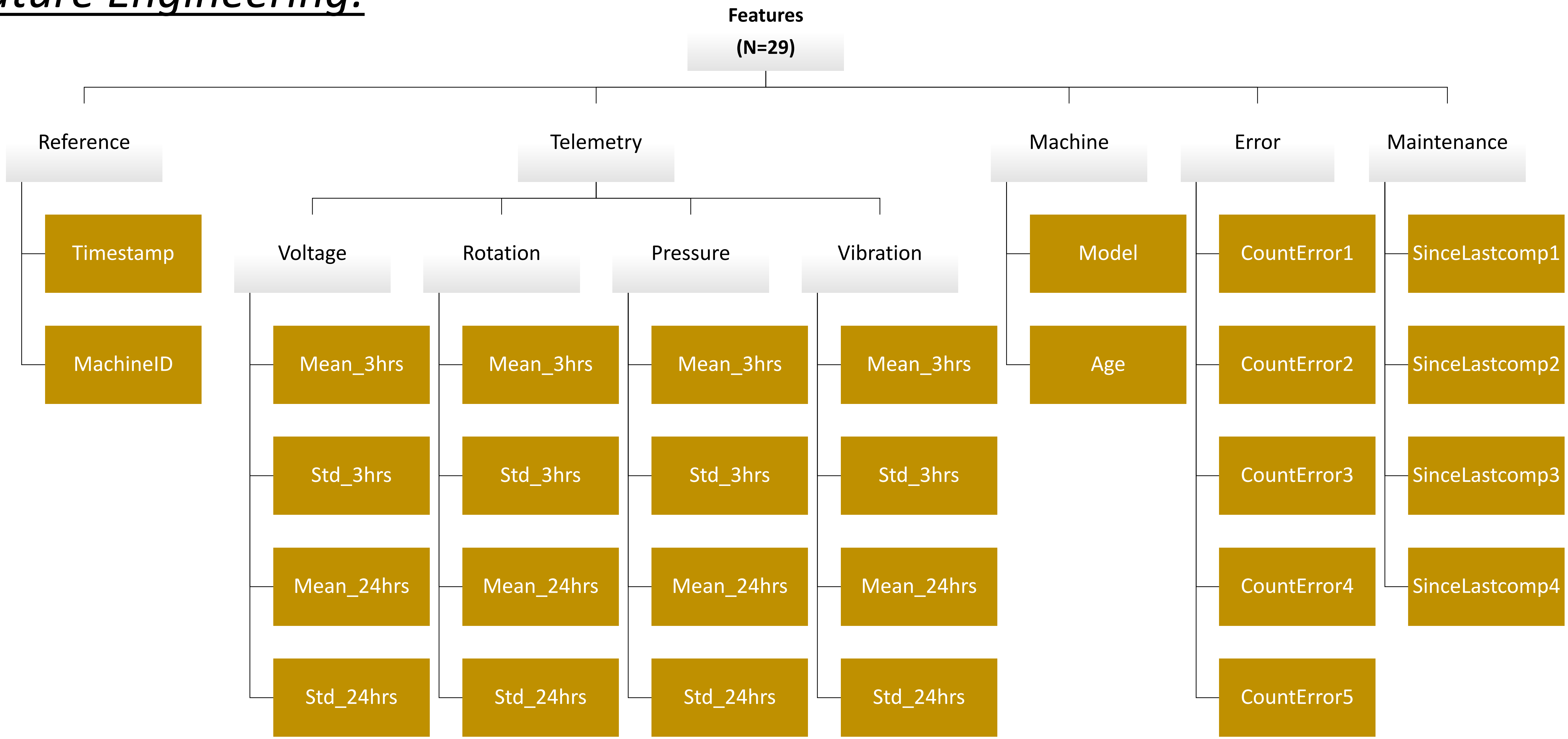
B. Feature Engineering:



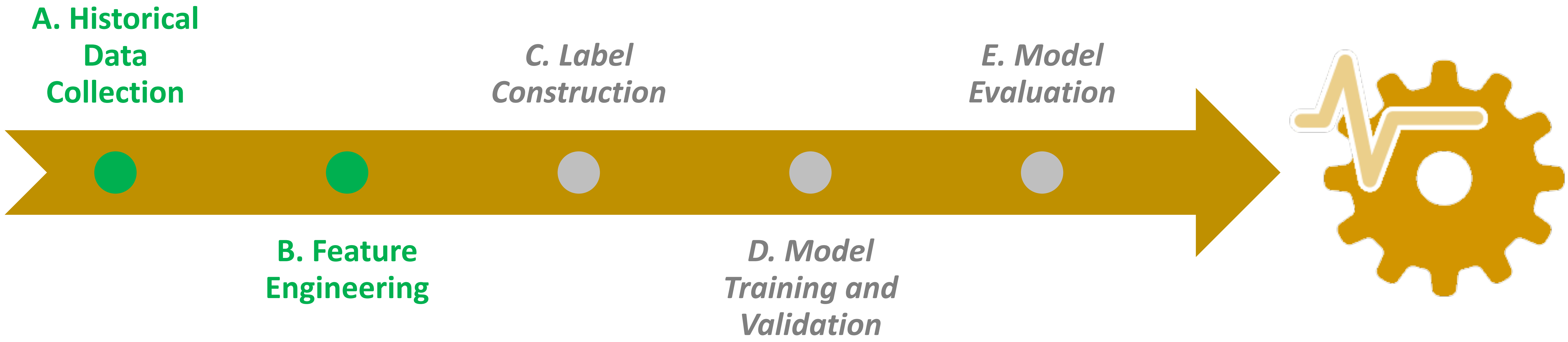


Example: Equipment Fault Diagnostics – Telemetry Data

B. Feature Engineering:

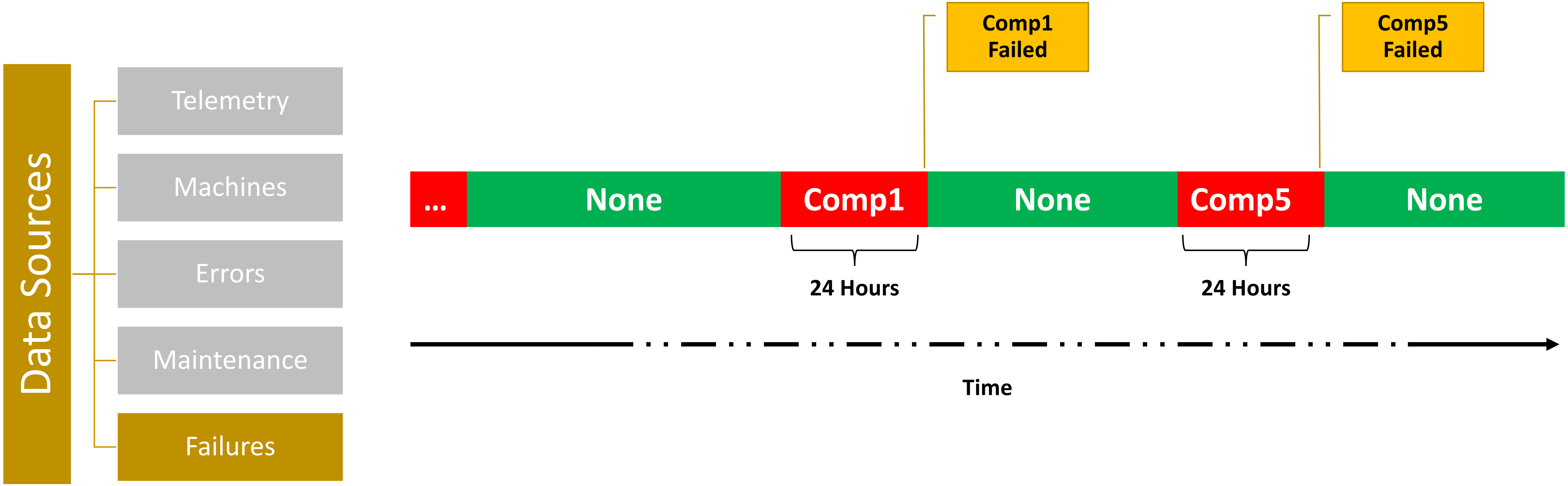


Example: Equipment Fault Diagnostics – Telemetry Data

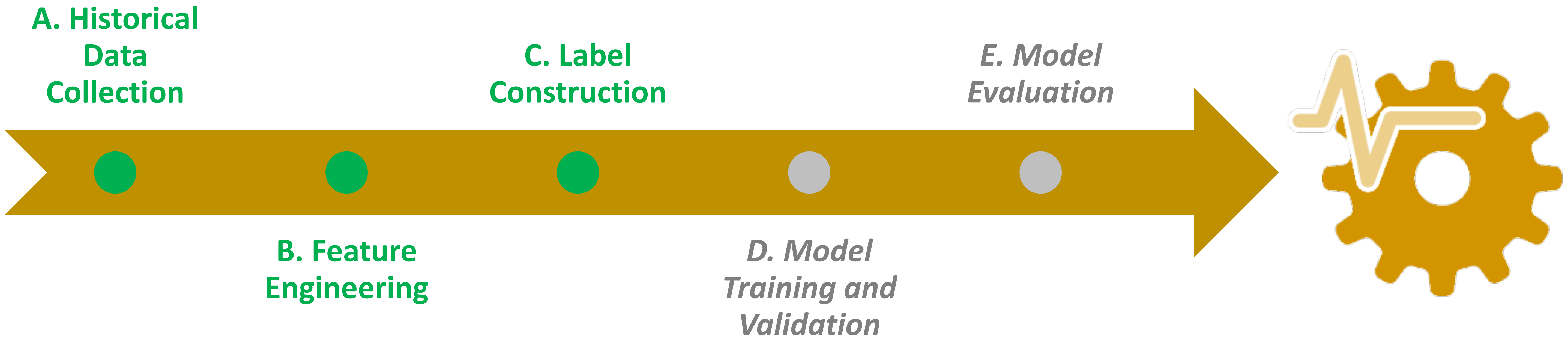


Example: Equipment Fault Diagnostics – Telemetry Data

C. Label Construction:

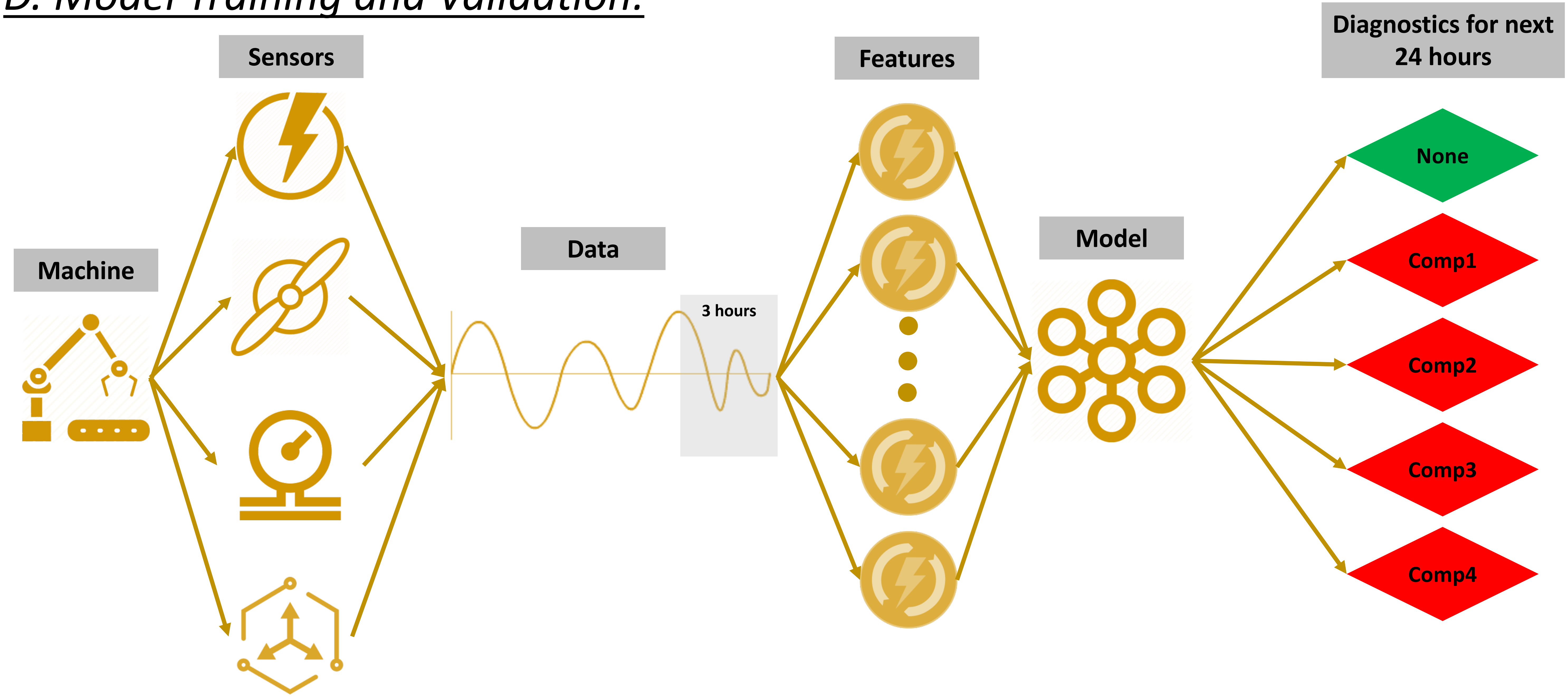


Example: Equipment Fault Diagnostics – Telemetry Data



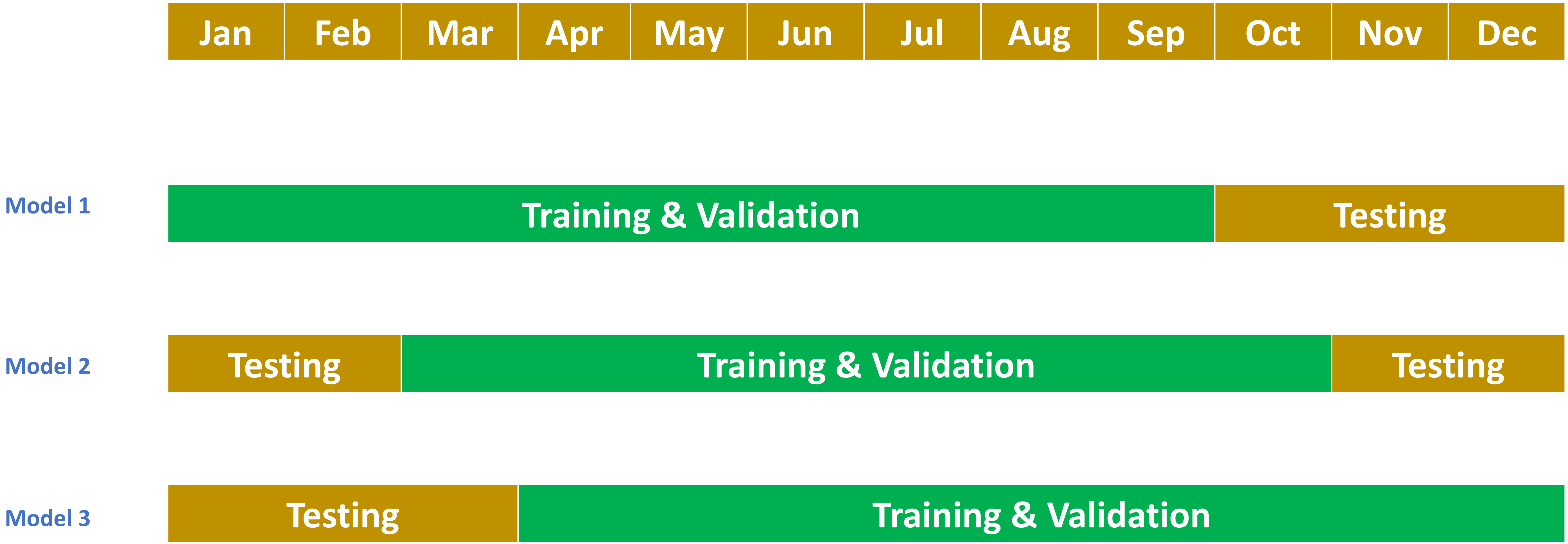
Example: Equipment Fault Diagnostics – Telemetry Data

D. Model Training and Validation:



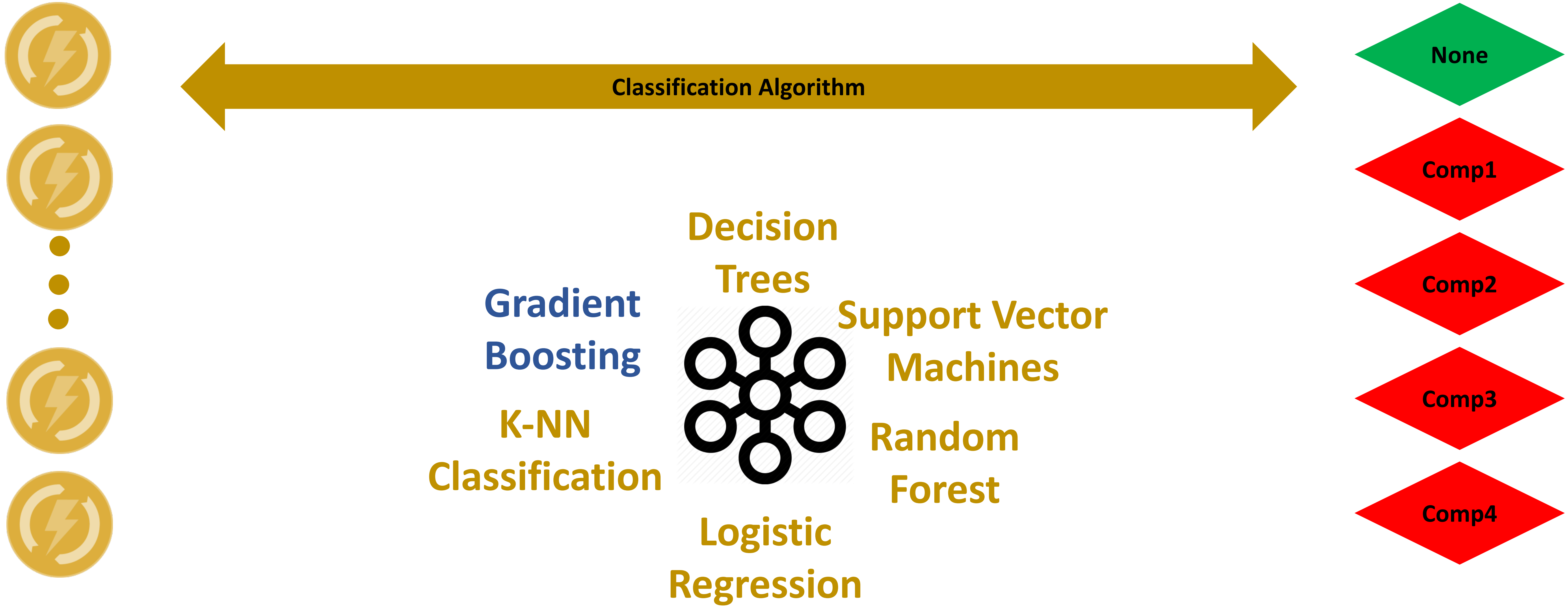
Example: Equipment Fault Diagnostics – Telemetry Data

D. Model Training and Validation:



Example: Equipment Fault Diagnostics – Telemetry Data

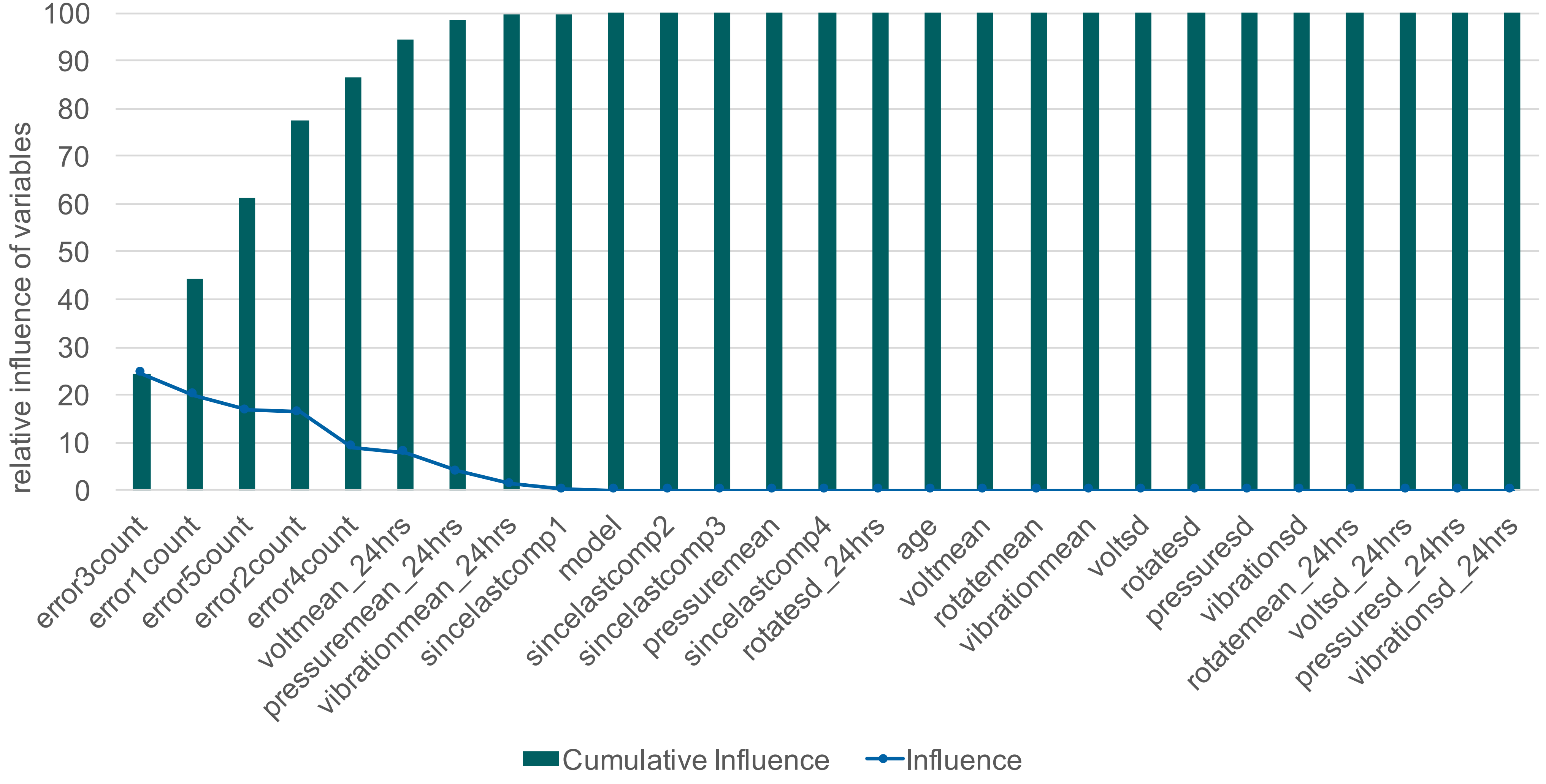
D. Model Training and Validation:



Example: Equipment Fault Diagnostics – Telemetry Data

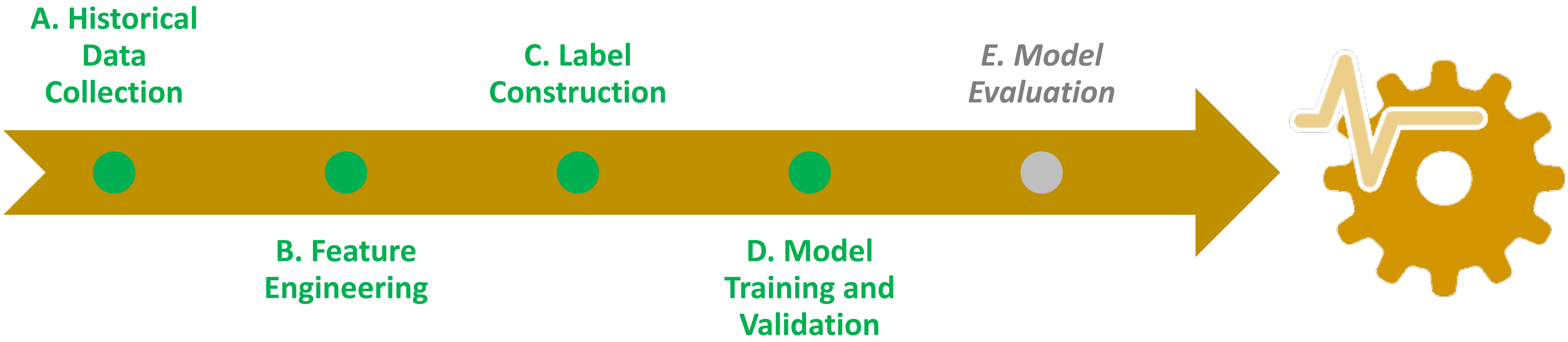
D. Model Training and Validation:

Model 3: Gradient Boosting



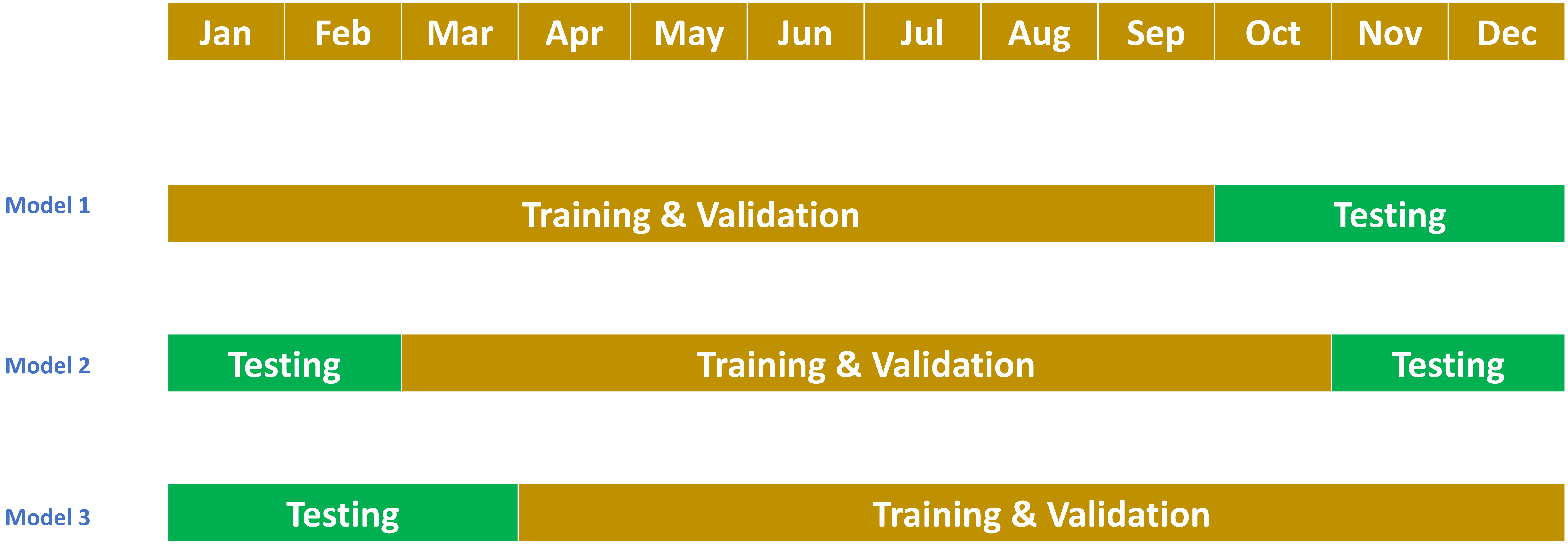
Variable	Influence
error3count	24.53
error1count	19.95
error5count	16.66
error2count	16.36
error4count	8.98
voltmean_24hrs	7.95
pressuremean_24hrs	3.96
vibrationmean_24hrs	1.34
sincelastcomp1	0.15
model	0.07
sincelastcomp2	0.02
sincelastcomp3	0.01
pressuremean	0.00
sincelastcomp4	0.00
rotatesd_24hrs	0.00
age	0.00
voltmean	0.00
rotatemean	0.00
vibrationmean	0.00
voltsd	0.00
rotatesd	0.00
pressuresd	0.00
vibrationsd	0.00
rotatemean_24hrs	0.00
voltsd_24hrs	0.00
pressuresd_24hrs	0.00
vibrationsd_24hrs	0.00

Example: Equipment Fault Diagnostics – Telemetry Data



Example: Equipment Fault Diagnostics – Telemetry Data

E. Model Evaluation:



Example: Equipment Fault Diagnostics – Telemetry Data

E. Model Evaluation:

		Predicted	
		Faulty	Healthy
Actual	Faulty	True Positive	False Negative (Missing Alarm)
	Healthy	False Positive (False Alarm)	True Negative

$$\text{Recall} = \frac{tp}{tp + fn}$$

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{True negative rate} = \frac{tn}{tn + fp}$$

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn}$$

Example: Equipment Fault Diagnostics – Telemetry Data

E. Model Evaluation:

		Predicted (model1)				
		comp1	comp2	comp3	comp4	none
Actual (model1)	comp1	5,684	33	96	63	4
	comp2	295	7,829	179	81	0
	comp3	0	5	3,203	0	0
	comp4	81	39	64	3,688	0
	none	66	5	0	0	1,203,521

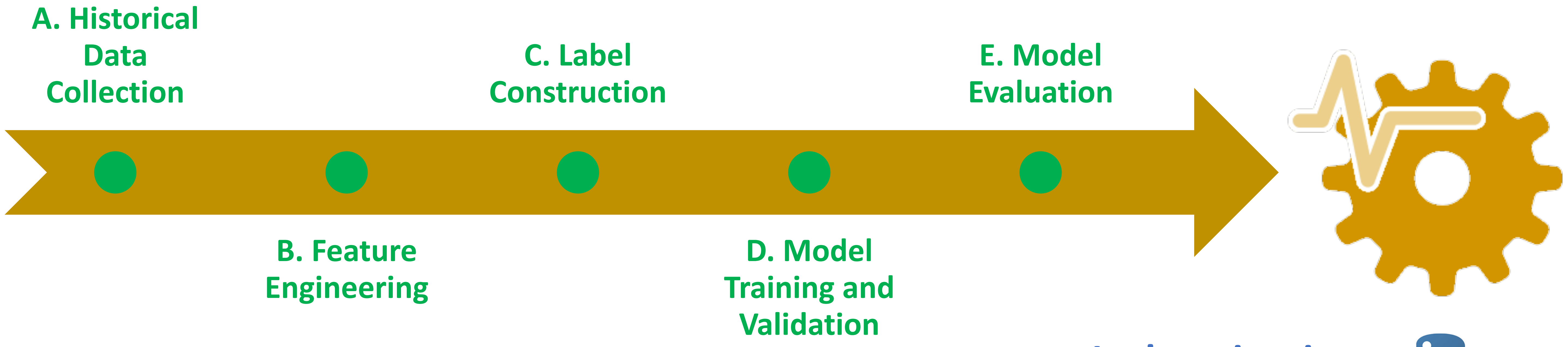
Missing Alarm = 4

False Alarm = 69

Class	comp1	comp2	comp3	comp4	none
Accuracy	0.99917				
Precision	0.92784	0.98963	0.90429	0.96242	0.99999
Recall	0.96666	0.93380	0.99844	0.95247	0.99994
F1	0.94685	0.96090	0.94903	0.95742	0.99996

Failure	Model1	Model2	Model3
comp1	**0.9667**	0.9466	0.9485
comp2	0.9338	**0.9714**	0.9509
comp3	**0.9984**	0.9104	0.9412
comp4	0.9525	0.9485	**0.9598**
none	0.9999	0.9999	**1.0000**

Example: Equipment Fault Diagnostics – Telemetry Data



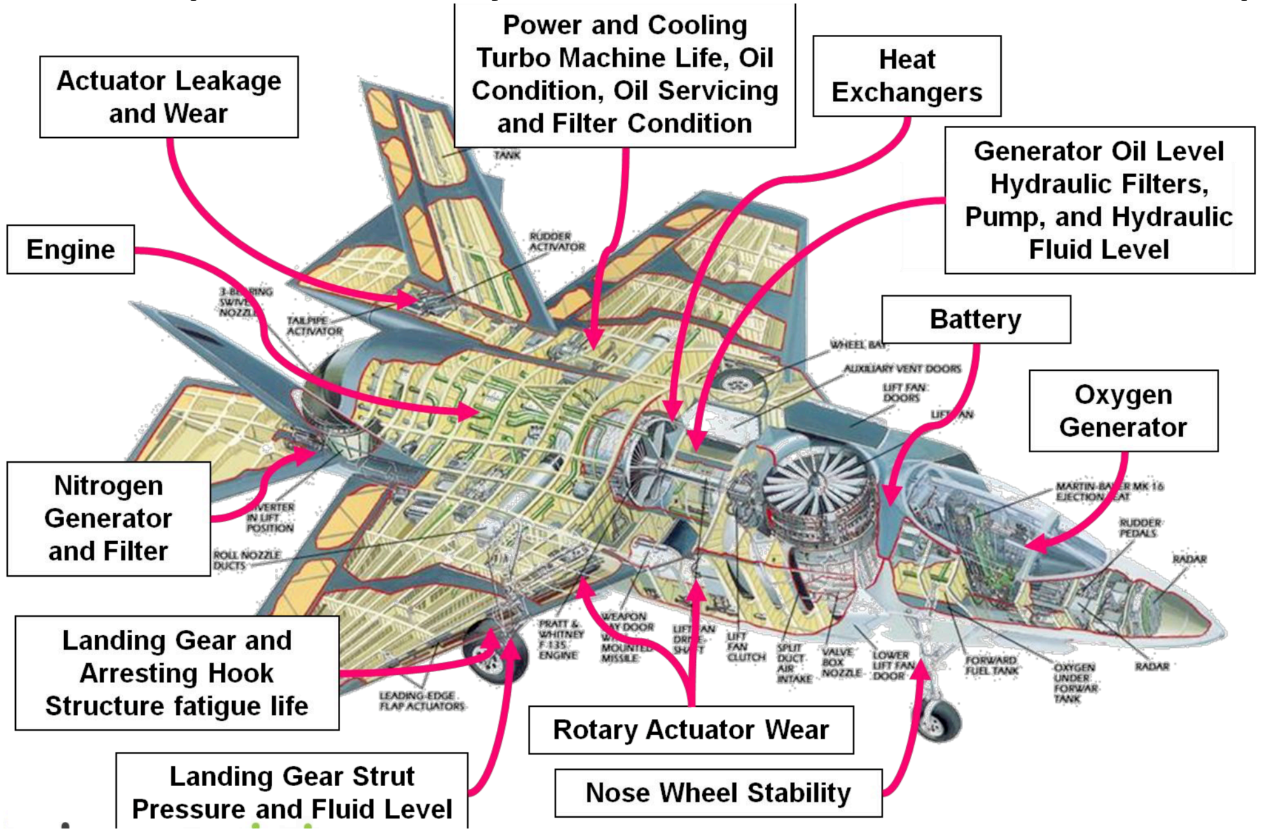
Let’s review it together

FAULT PROGNOSTICS

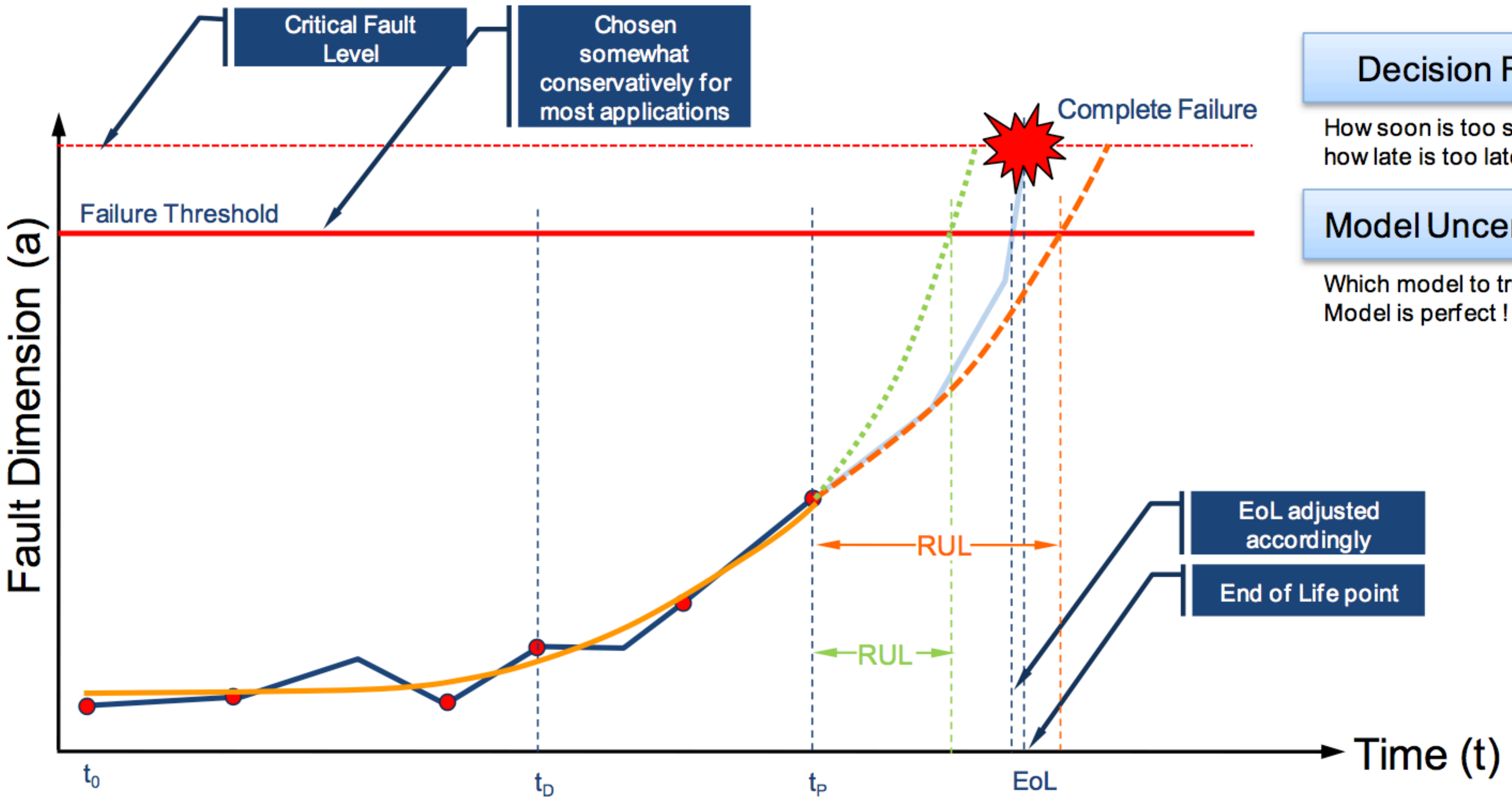
WHAT IS PROGNOSTICS?

- PHM Community – “Estimation of the Remaining Useful Life of a component”
- Remaining Useful Life (RUL) – The amount of time a component can be expected to continue operating (Not necessarily to failure!) within its given specifications.
    - Dependent on future operating conditions (input commands, environment, and loads)
- Time of Failure (TOF): the time a component is expected to fail (no longer meet its design specifications).
- Probability of Failure (POF): the failure probability distribution of the component.

F-35 Prognostics



PROGNOSTICS FRAMEWORK



Decision Risk

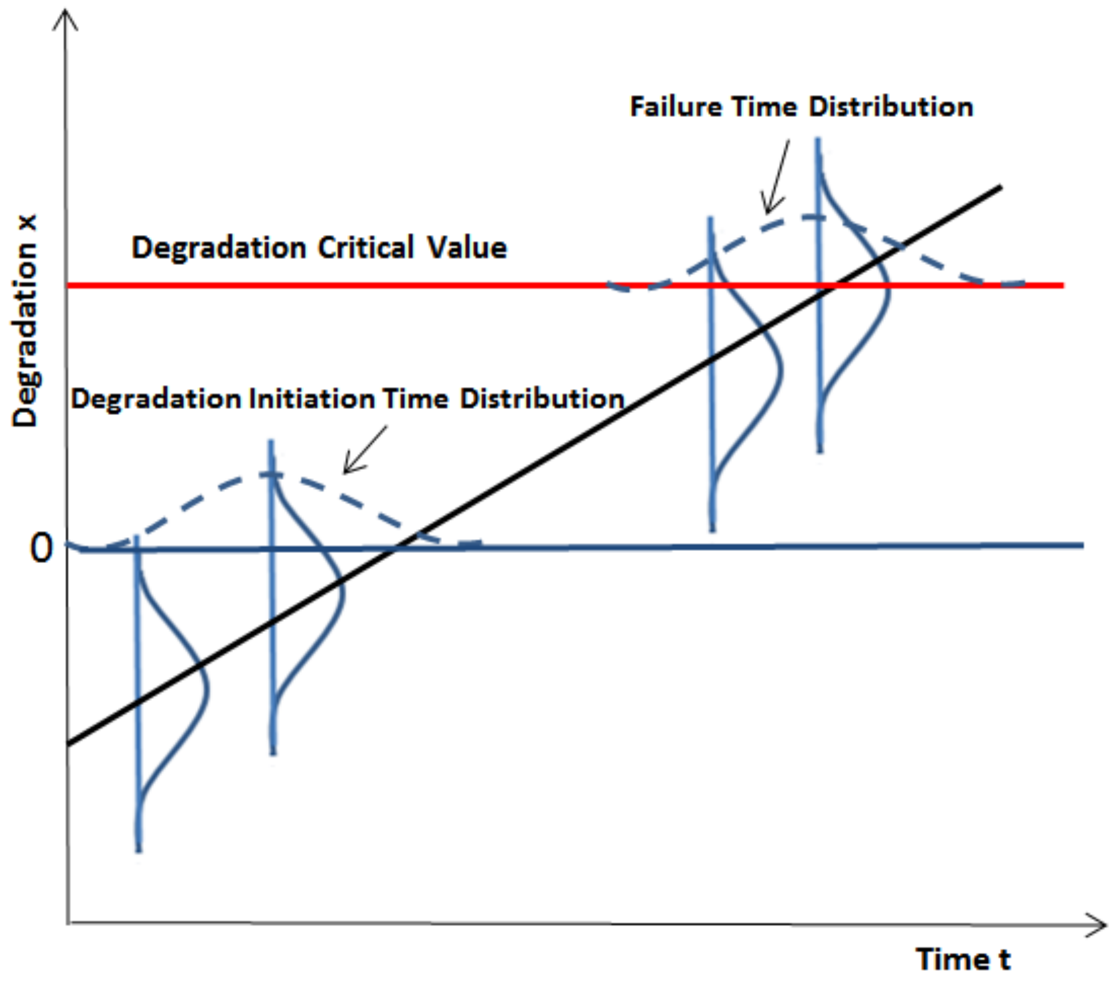
How soon is too soon and how late is too late?

Model Uncertainty

Which model to trust? No Model is perfect !

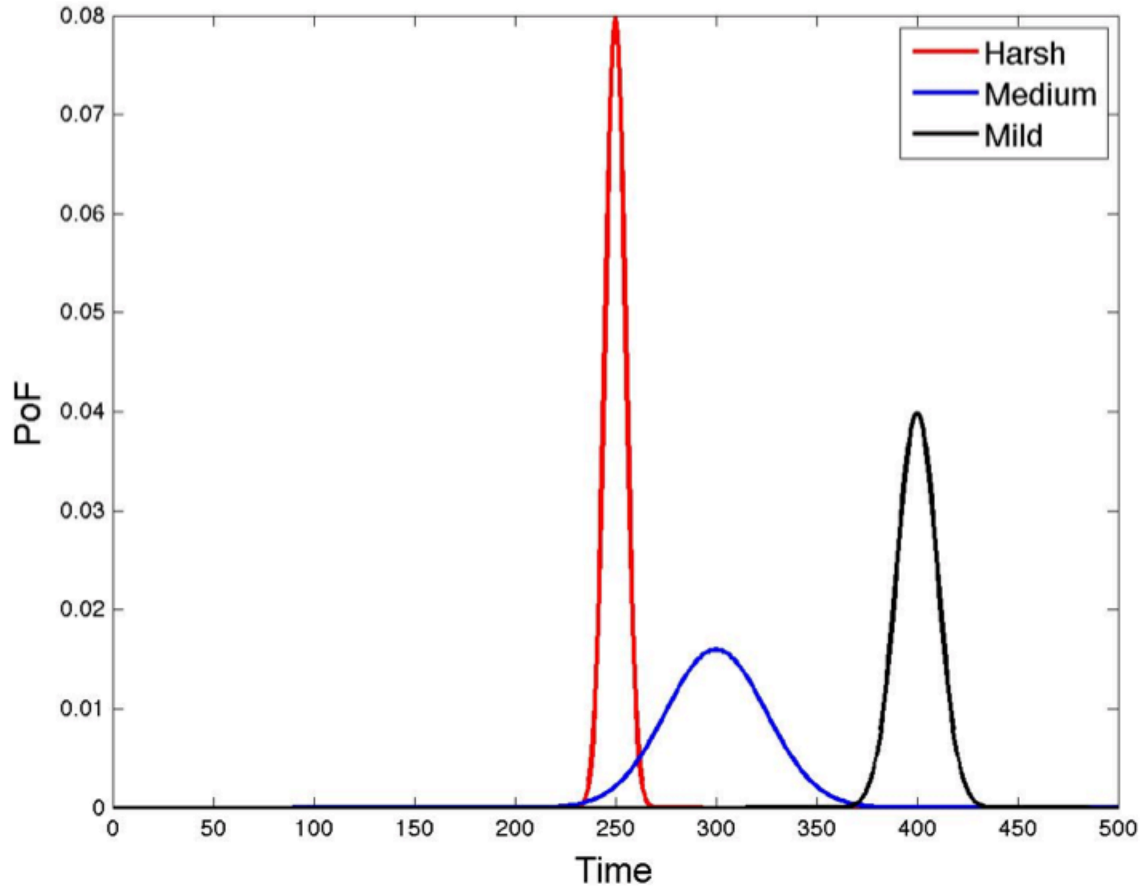
TYPES OF PROGNOSTICS METHODS

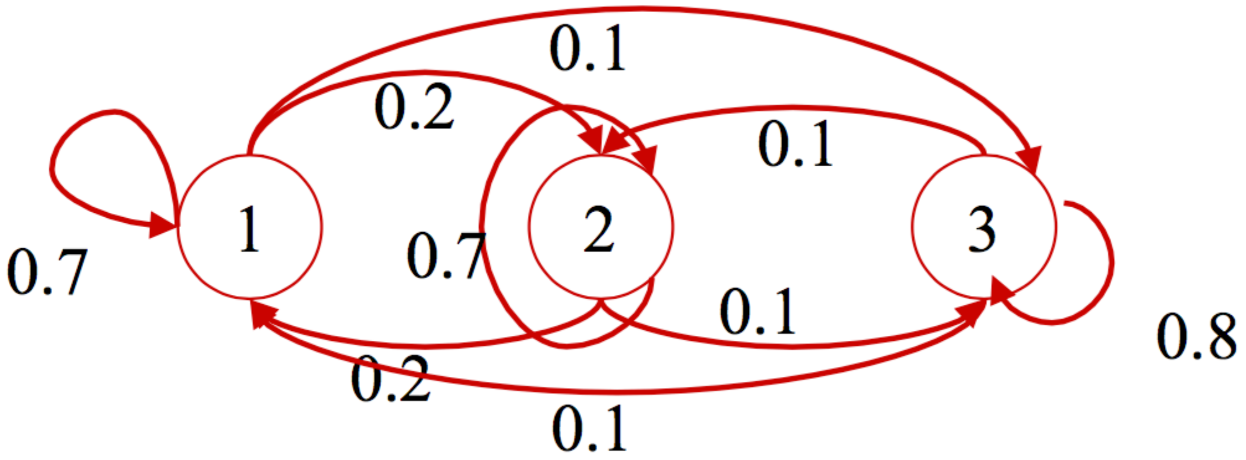
- Type I: Reliability Data-based
  - Use population based statistical model
  - These methods consider historical time to failure data which are used to model the failure distribution. They estimate the life of a typical component under nominal usage conditions.
  - Ex: Weibull Analysis



TYPES OF PROGNOSTICS METHODS

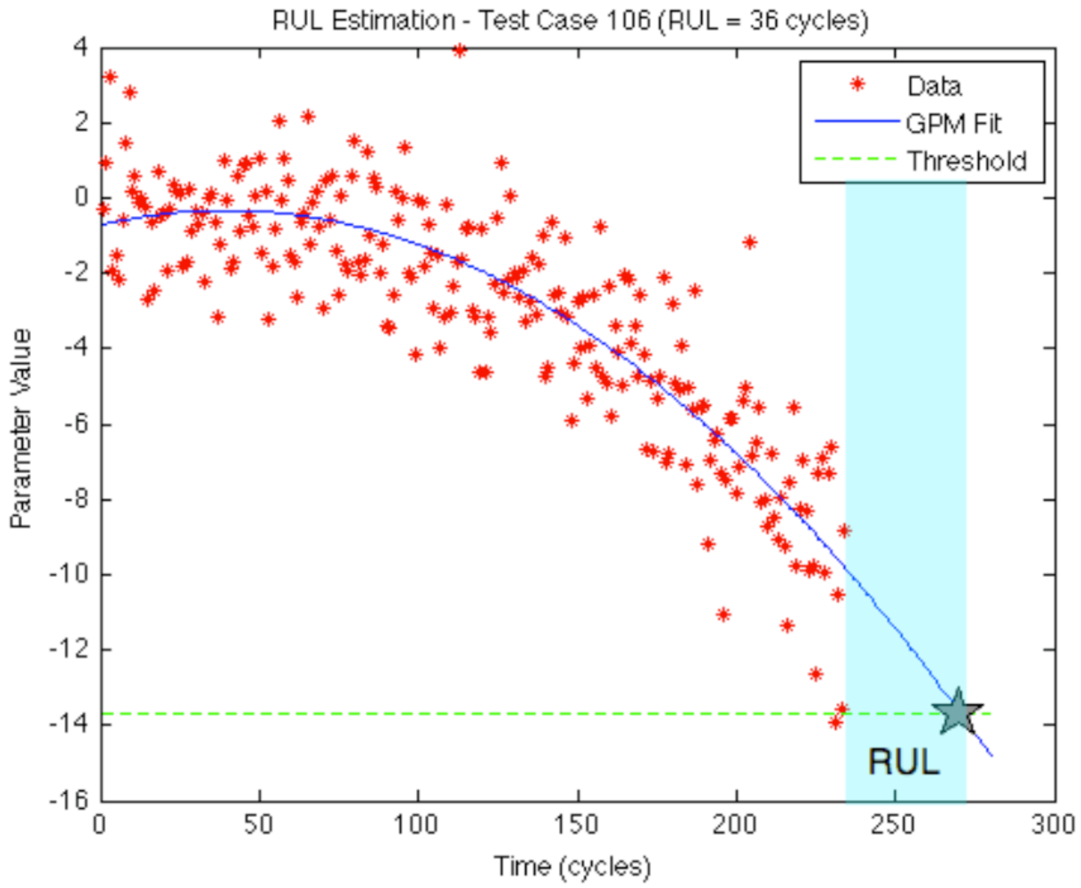
- Type II: Stress-based
  - Use population based fault growth model – learned from accumulated knowledge
  - These methods also consider the environmental stresses (temperature, load, vibration, etc.) on the component. They estimate the life of an average component under specific usage conditions.
  - Ex: Proportional Hazards Model, Markov Chain Models



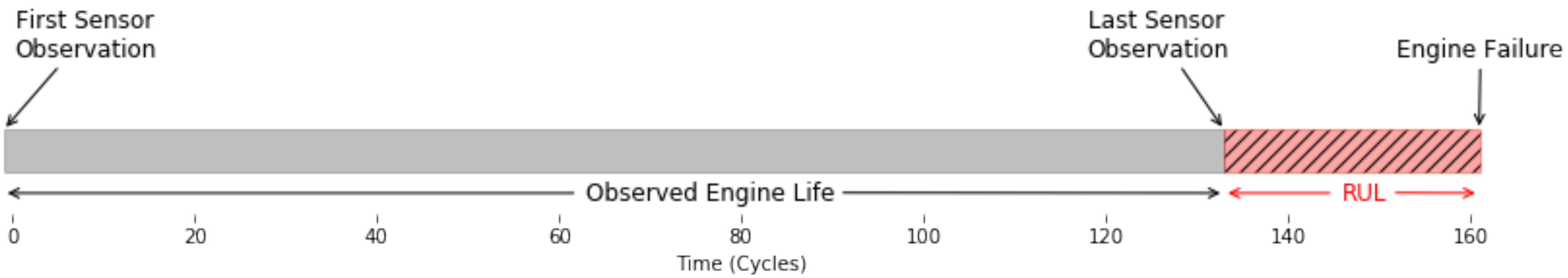


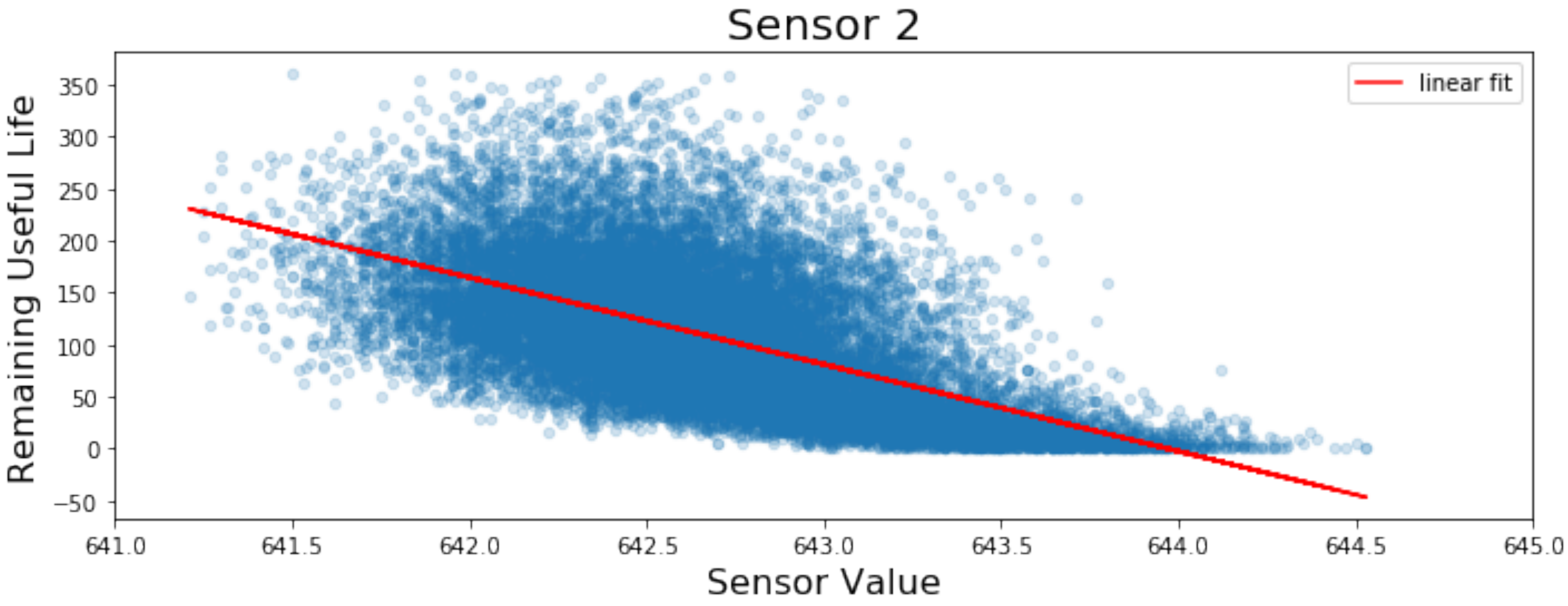
TYPES OF PROGNOSTICS METHODS

- Type III: Condition-based
  - Individual component based data-driven model
  - Degradation must be related to a measurable parameter such as tread depth or bearing vibration level or temperature or inferred from available measurements
  - Ex: Cumulative Damage Model, Filtering and State Estimation

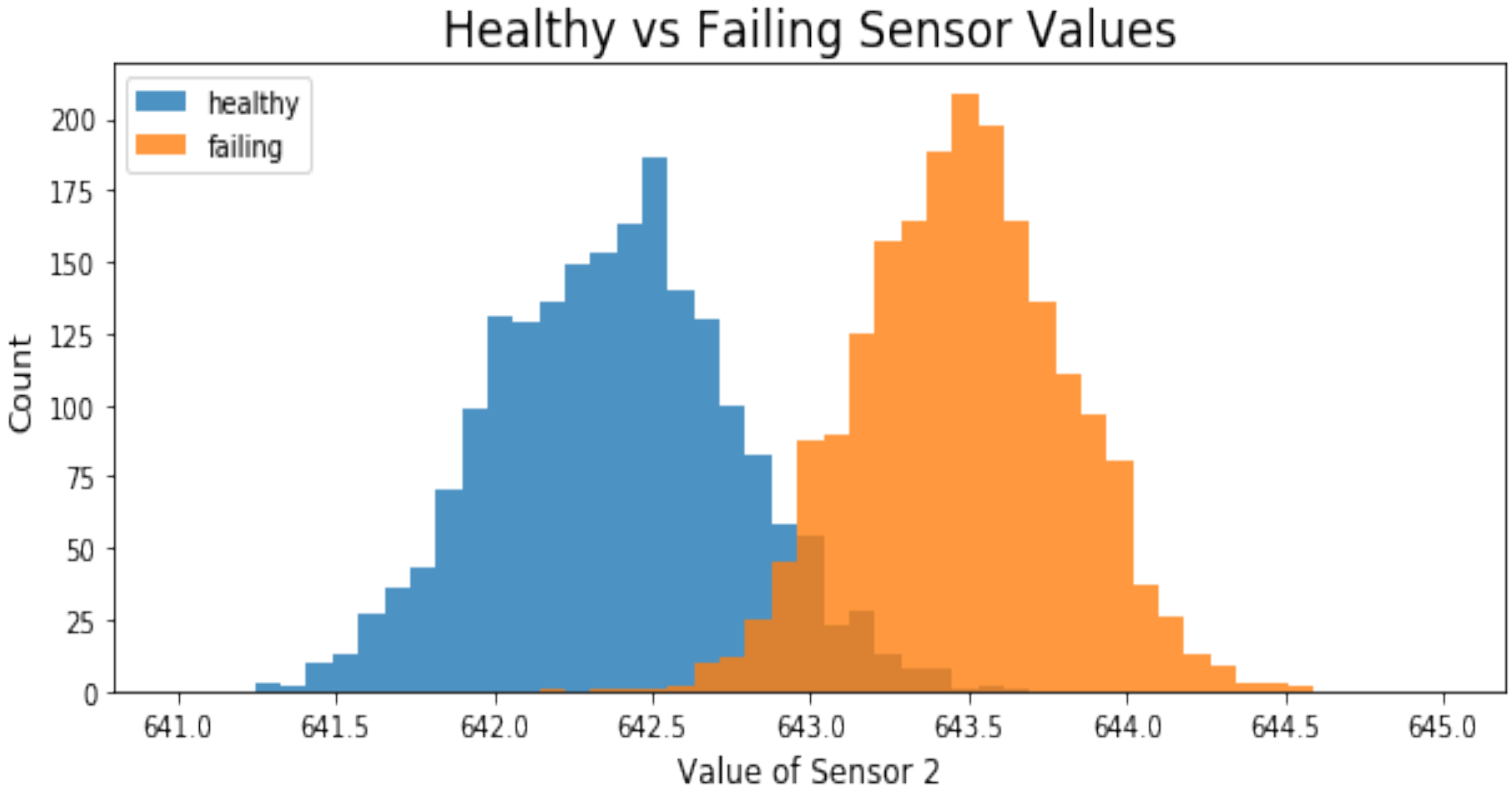


PROGNOSTICS DATA





PROGNOSTICS DATA

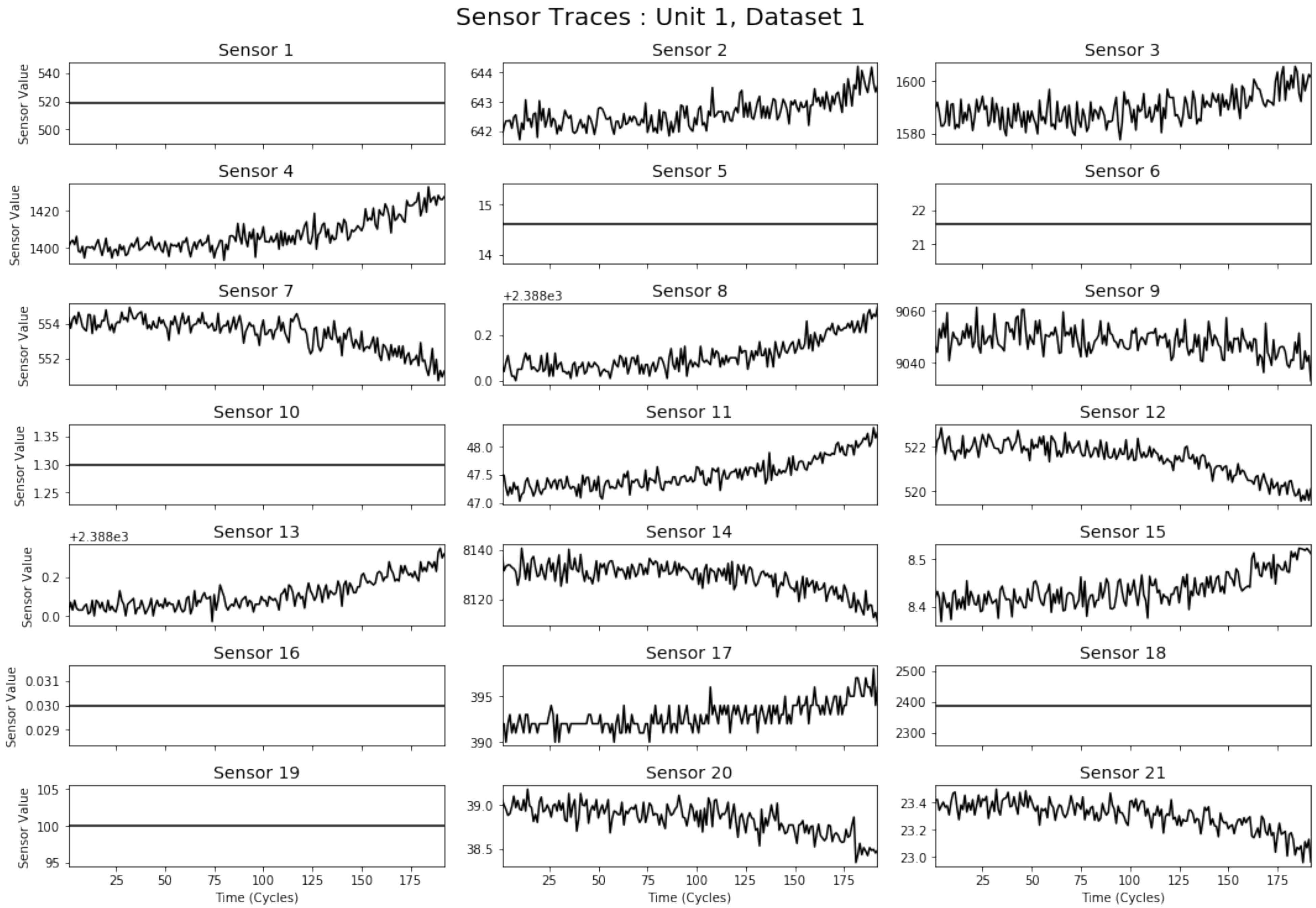


NASA DATA PROGNOSTICS EXAMPLE

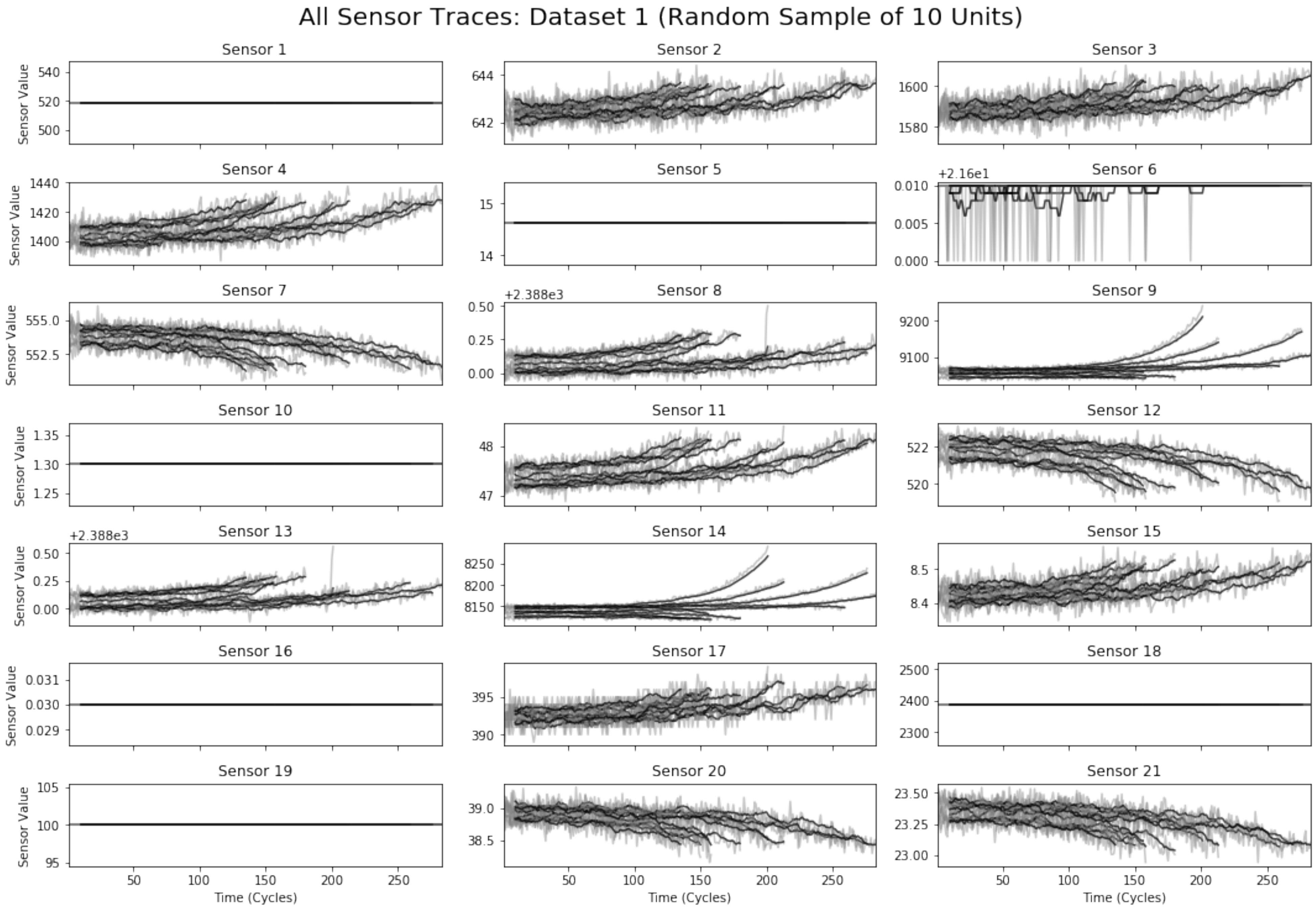
Column	Description
dataset_id	id of the original dataset of this instance
unit_id	id of engine (unique in each dataset)
cycle	number of operational cycles since beginning of engine operation
setting 1	value of operational setting 1
setting 2	value of operational setting 2
setting 3	value of operational setting 3
sensor 1	value of sensor 1
...	...
sensor 21	value of sensor 21

Let's explore an open data set that poses a related problem. The Turbofan Engine Degradation Simulation data set was released in 2008 by the Prognostics Center of Excellence at NASA's Ames research center. It consists of sensor readings from a fleet of simulated aircraft gas turbine engines, recorded as multiple multivariate time series.

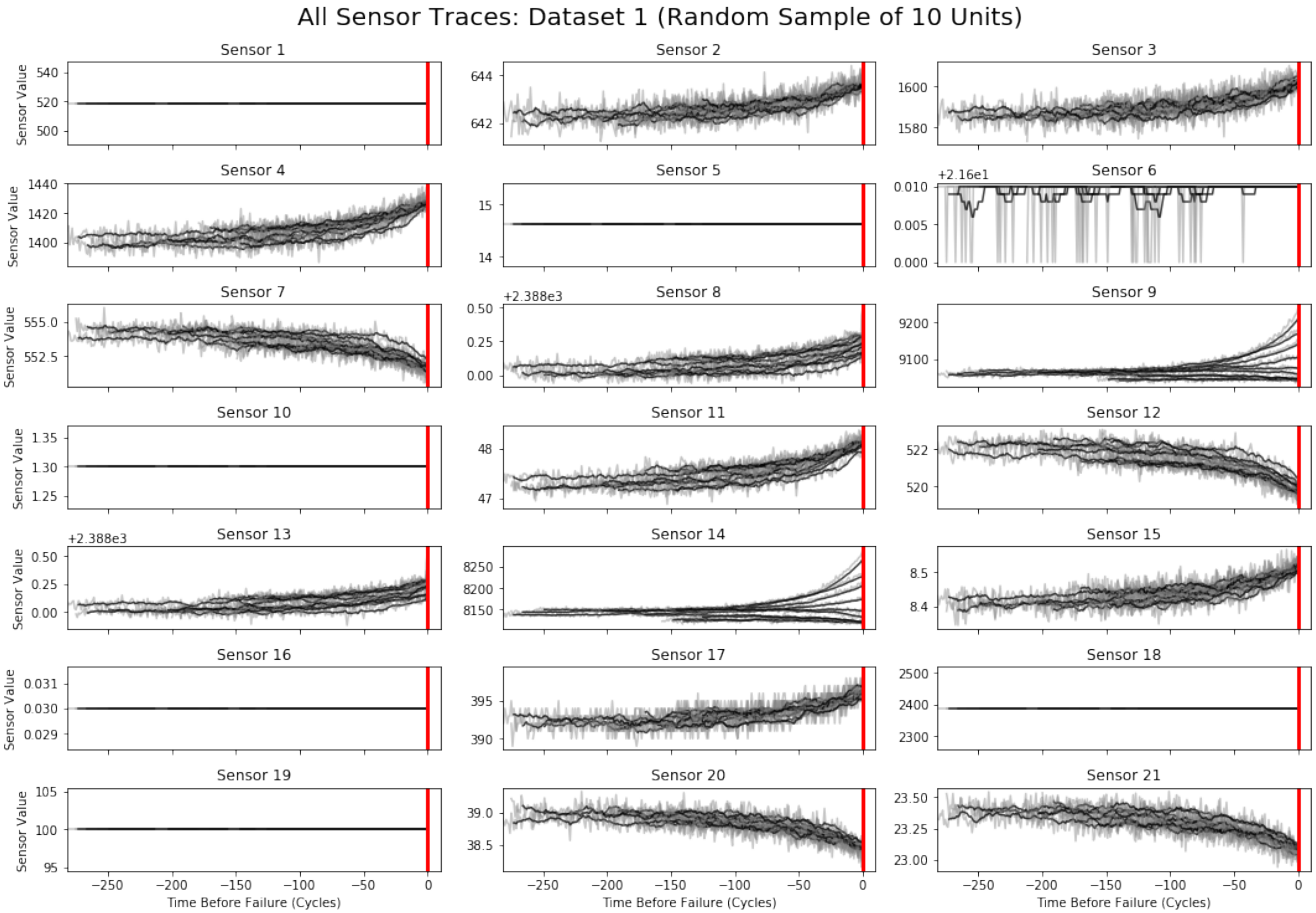
NASA DATA PROGNOSTICS EXAMPLE



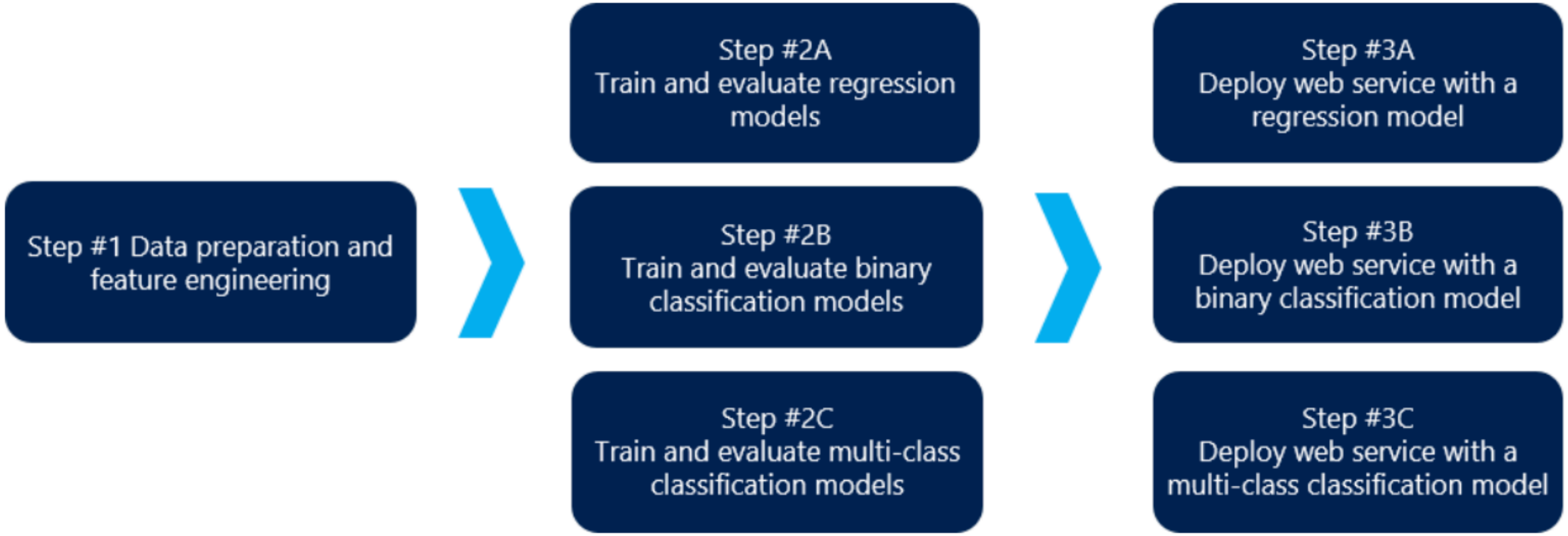
NASA DATA PROGNOSTICS EXAMPLE



NASA DATA PROGNOSTICS EXAMPLE



NASA DATA PROGNOSTICS EXAMPLE



- Step 1: Data preparation and feature engineering
- Step 2: Train and evaluate model
- Step 3: Deploy as web service

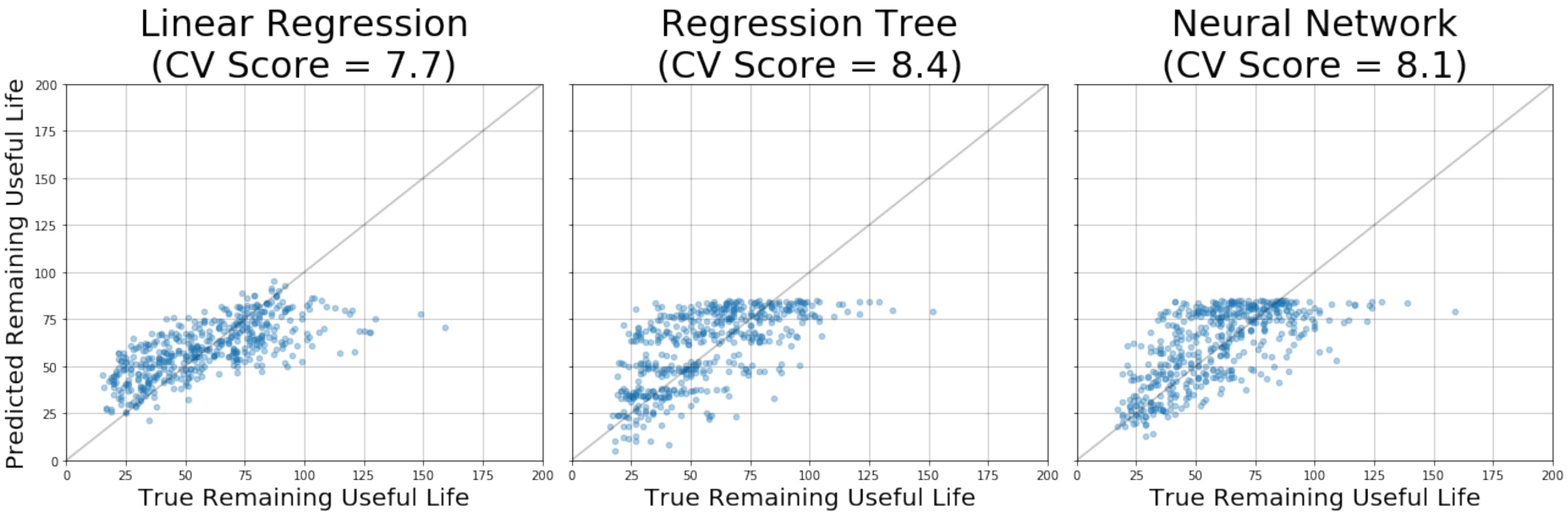
NASA DATA PROGNOSTICS EXAMPLE

id	cycle	...	RUL	label1	label2
1	1		191	0	0
1	2		190	0	0
1	3		189	0	0
1	4		188	0	0
...			...	...	...
1	160		32	0	0
1	161		31	0	0
1	162		30	1	1
1	163		29	1	1
1	164		28	1	1
1	165		27	1	1
1	166		26	1	1
1	167		25	1	1
1	168		24	1	1
1	169		23	1	1
1	170		22	1	1
1	171		21	1	1
1	172		20	1	1
1	173		19	1	1
1	174		18	1	1
1	175		17	1	1
1	176		16	1	1
1	177		15	1	2
1	178		14	1	2
1	179		13	1	2
1	180		12	1	2
1	181		11	1	2
1	182		10	1	2
1	183		9	1	2
1	184		8	1	2
1	185		7	1	2
1	186		6	1	2
1	187		5	1	2
1	188		4	1	2
1	189		3	1	2
1	190		2	1	2
1	191		1	1	2
1	192		0	1	2

w0 = 15
w1 = 30

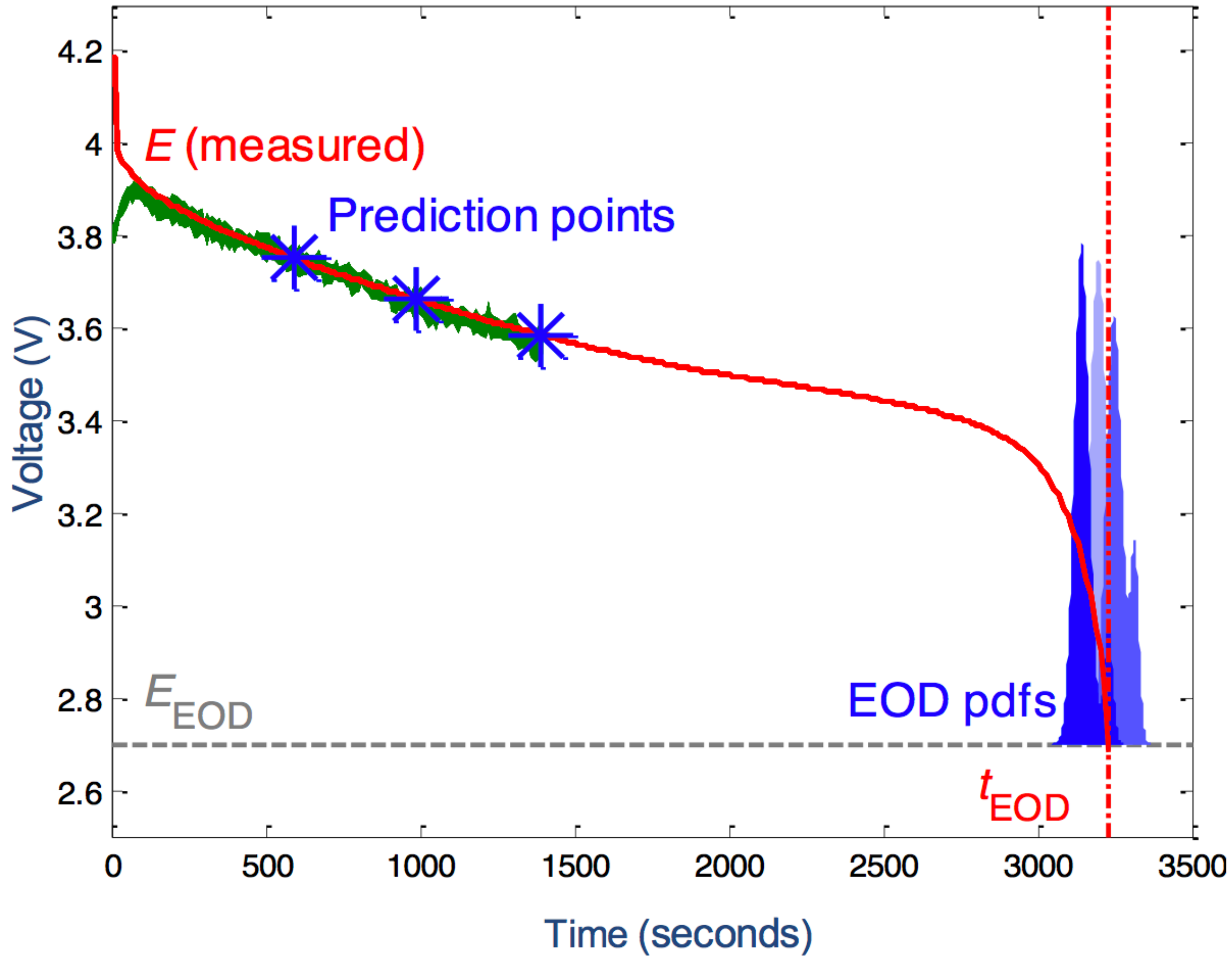
w1

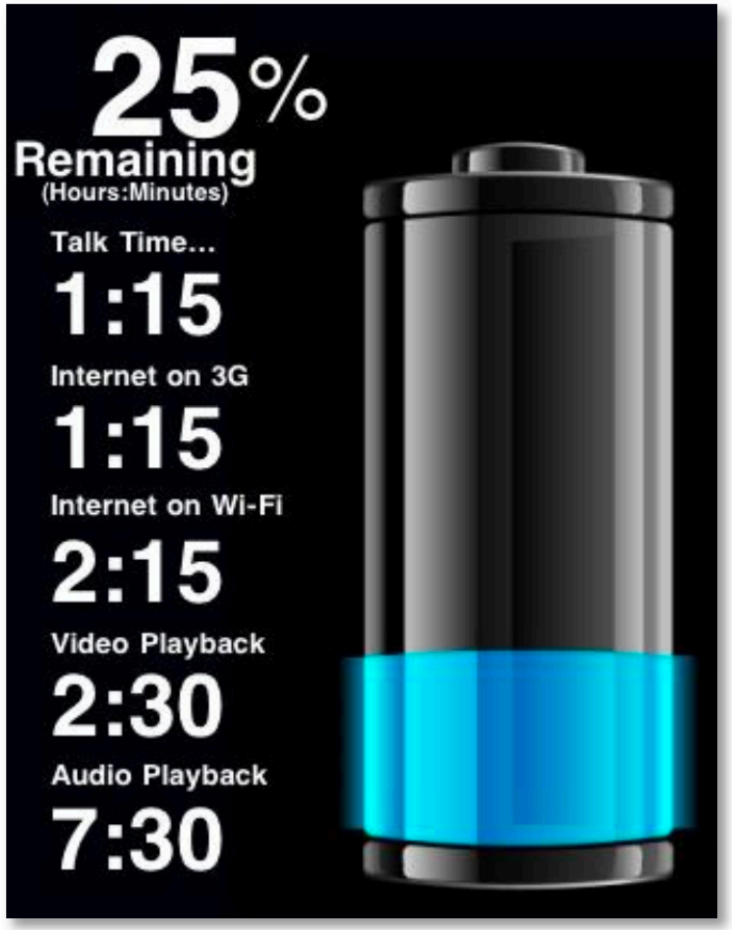
w0



Predicting Battery Discharge

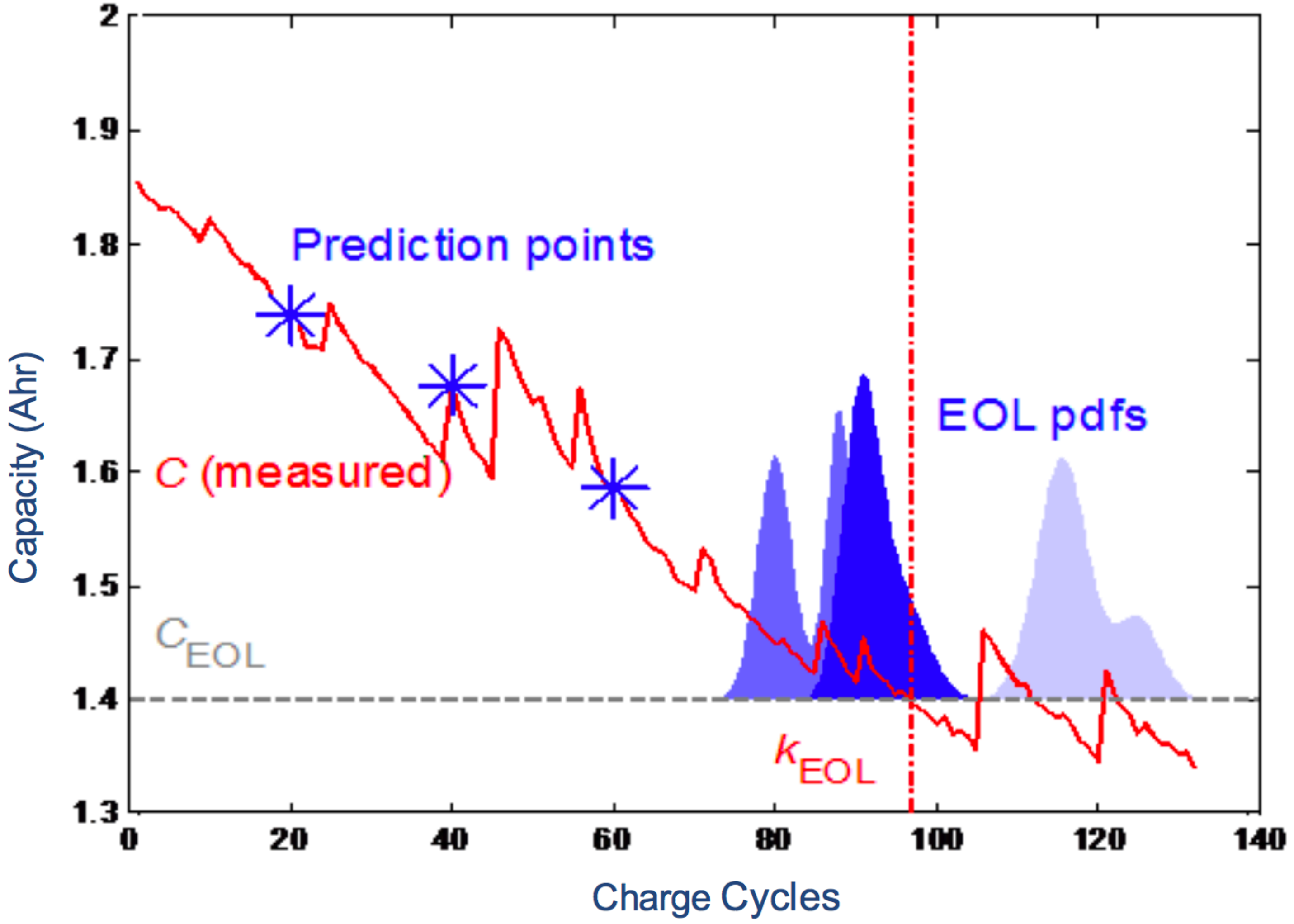
- Objective: Predict when the battery voltage will dip below 2.7 volts
- Example: when to recharge laptop or cell phone batteries





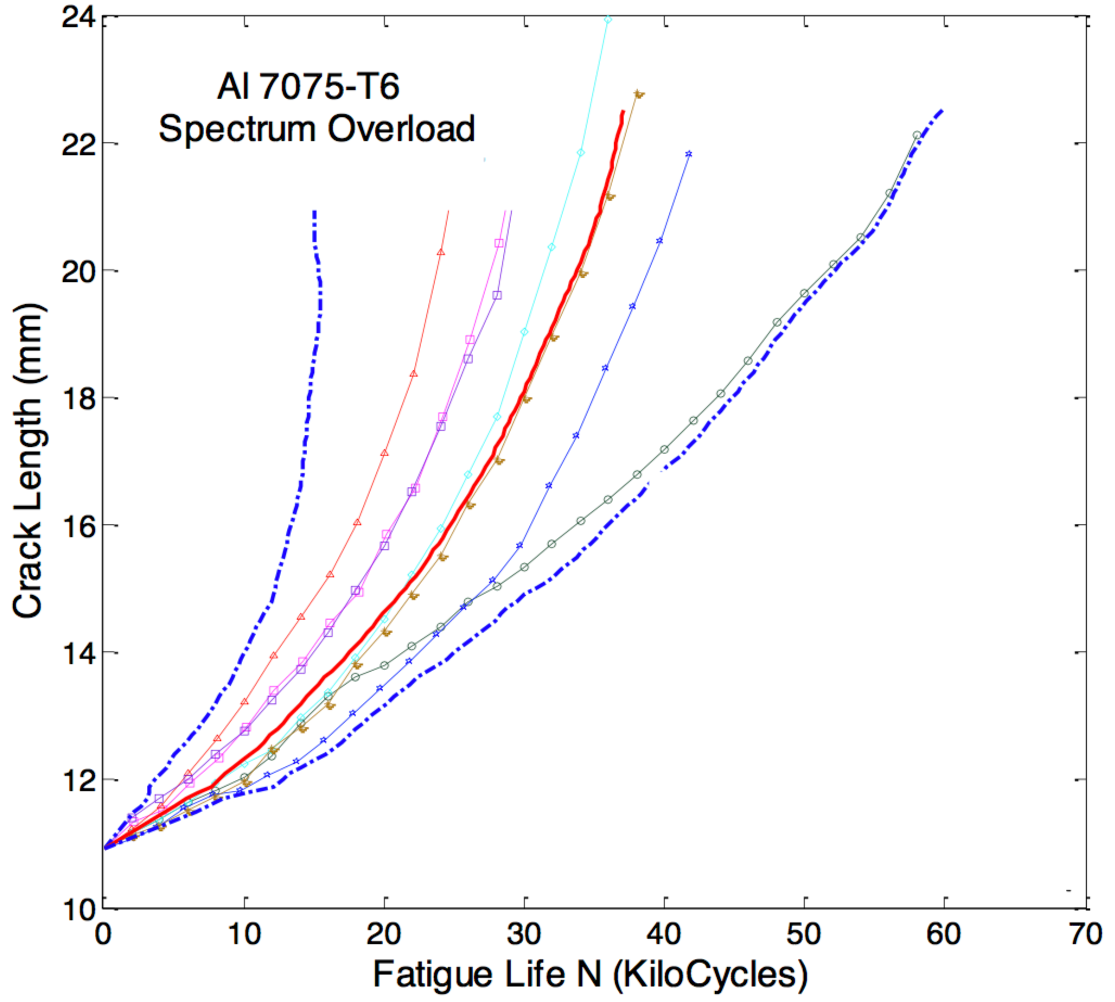
Predicting Battery Capacity – Long Term

- Objective: Predict when Li-ion battery capacity will fade by 30% indicating life (EOL)
- Example: when to replace batteries



Predicting crack size in metallic structures

- Objective: Predict when the crack size will exceed a critical length
- Example: aircraft structures, bridges, buildings, etc.



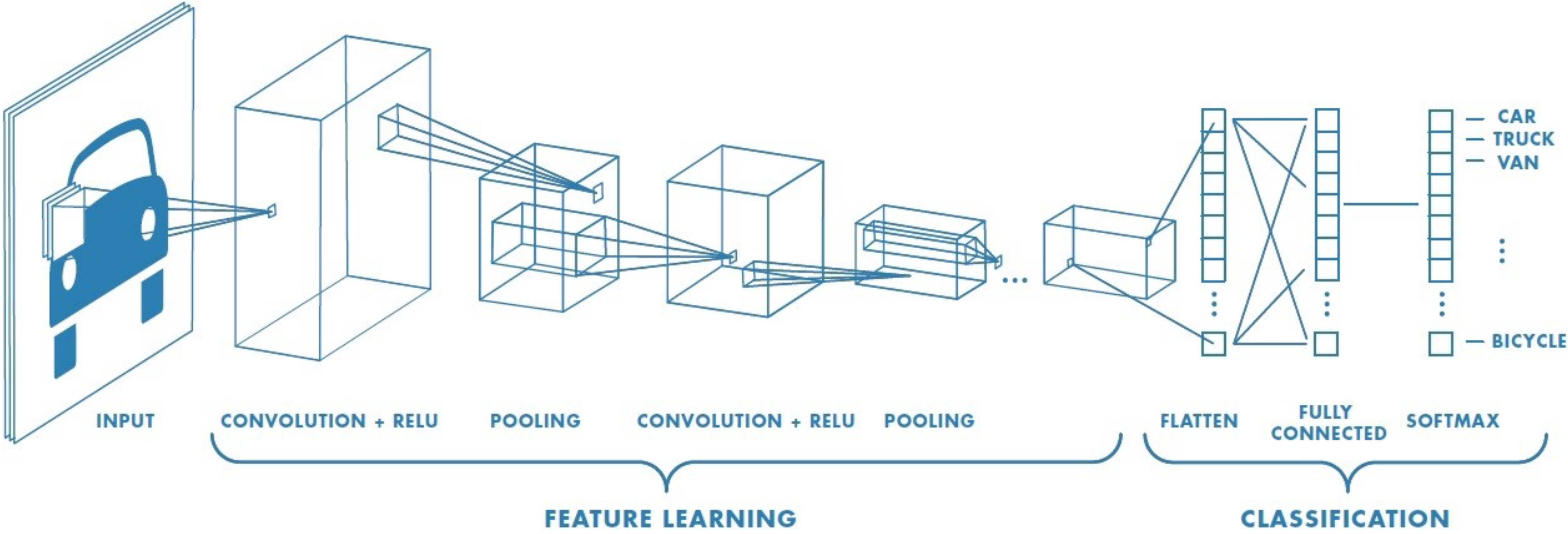
Deep Learning

WHAT IS DEEP LEARNING?

Set of methods to learn very flexible differentiable functions that seem to work well for most real life data.

An Artificial Neural Network (ANN) is a computational model that is inspired by the way biological neural networks in the human brain process information.

HOW THAT WORKS?



Let’s see it together:

https://teachablemachine.withgoogle.com

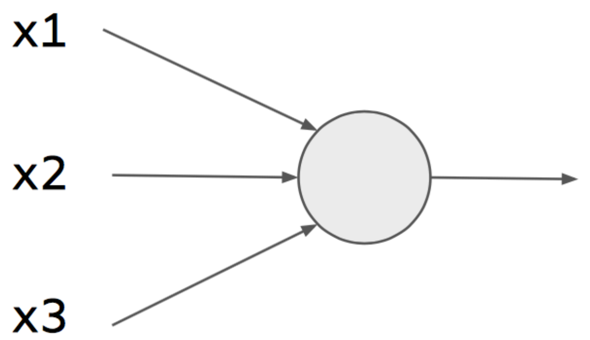
HOW IT WORKS

Input

Music: Yes/No

Friends: Yes/No

Weather: Good/bad



Output

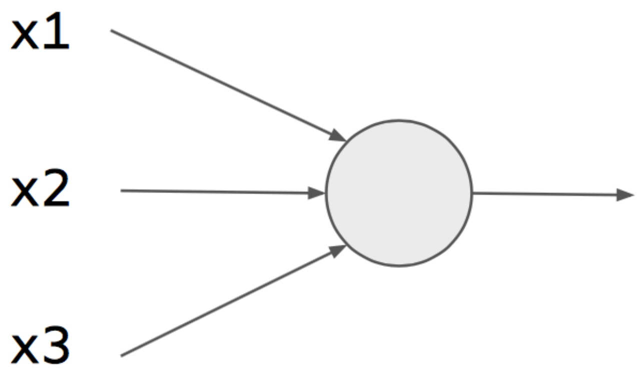
Go out: Yes/No

HOW IT WORKS

Music: 0

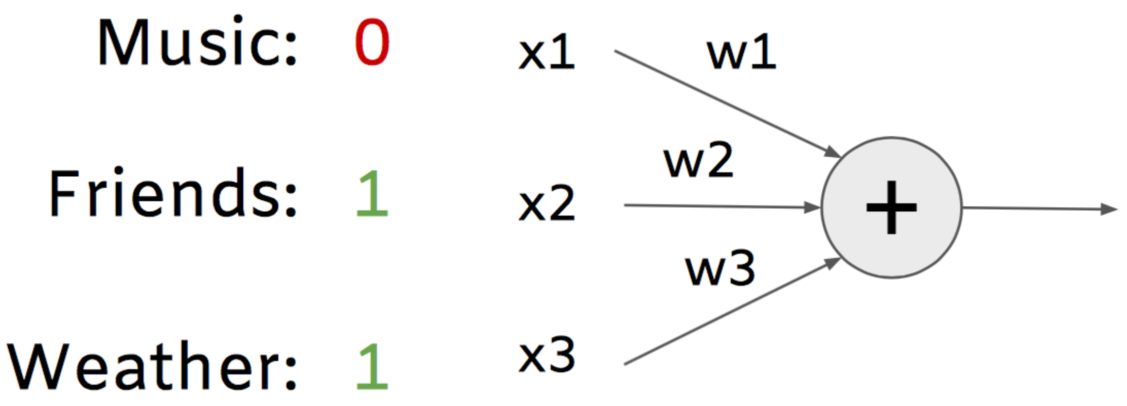
Friends: 1

Weather: 1



Go out: 1

HOW IT WORKS



$= x1 * w1 + x2 * w2 + x3 * w3$

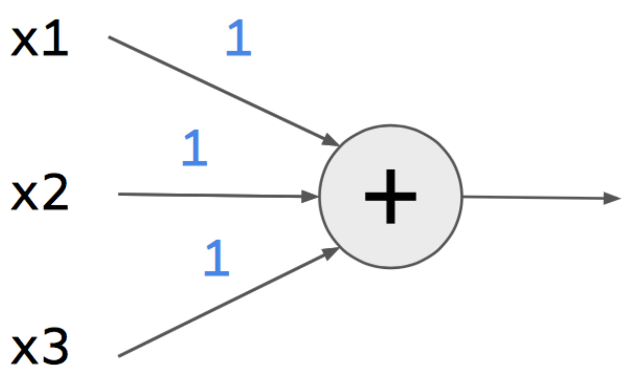
$= 0 * w1 + 1 * w2 + 1 * w3$

HOW IT WORKS

Music: 0

Friends: 1

Weather: 1

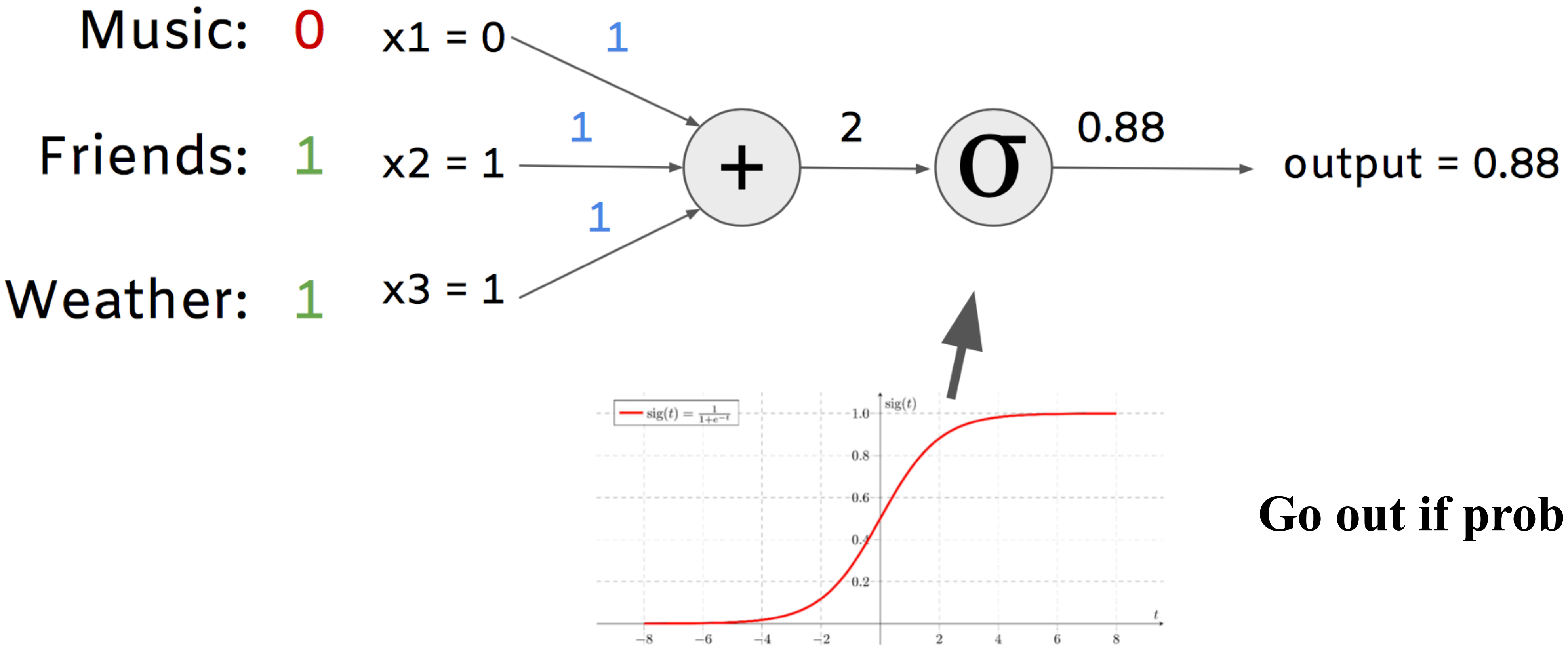


$$= x1 * w1 + x2 * w2 + x3 * w3$$

$$= 0 * 1 + 1 * 1 + 1 * 1$$

$$= 2$$

HOW IT WORKS



Go out if probability is > 0.5

HOW IT WORKS

We now have a neural net that can make simple decisions

How can we train the neural net, i.e. **adjust w1, w2, w3** so that the decisions it makes are "correct"?

HOW IT WORKS

Define a loss or cost, i.e. a way to tell the neural network how bad it is doing

Loss = (output - decision)^2

For example

Loss = (0.88 - 1)^2 = 0.014

HOW IT WORKS

Choose w1, w2, w3 so the loss is minimized
on the examples

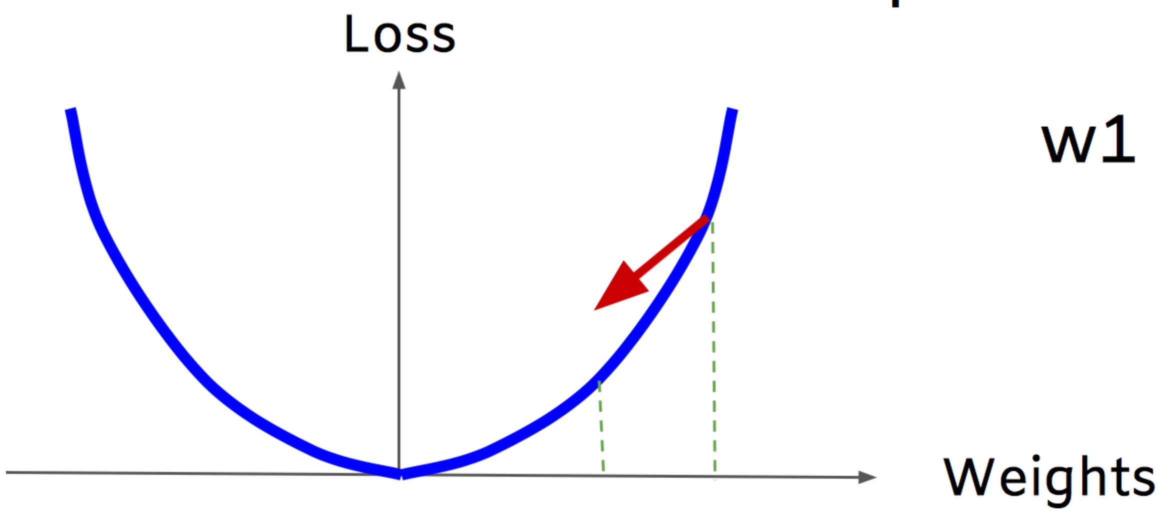
How to adjust w1, w2, w3?

HOW IT WORKS

Calculate **derivative of loss with respect to weights**

$$\frac{dL}{dw1}$$

Update weights to minimize loss



$$w1 \leftarrow w1 - a\,\frac{dL}{dw1}$$

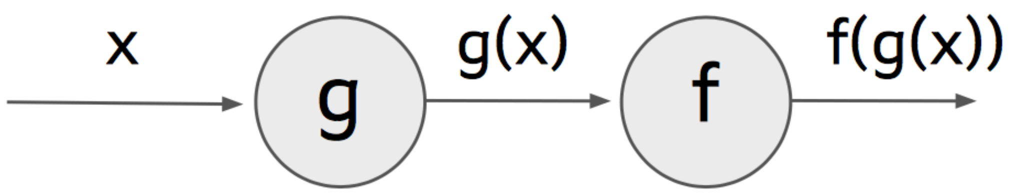
HOW IT WORKS

How to combine nodes?

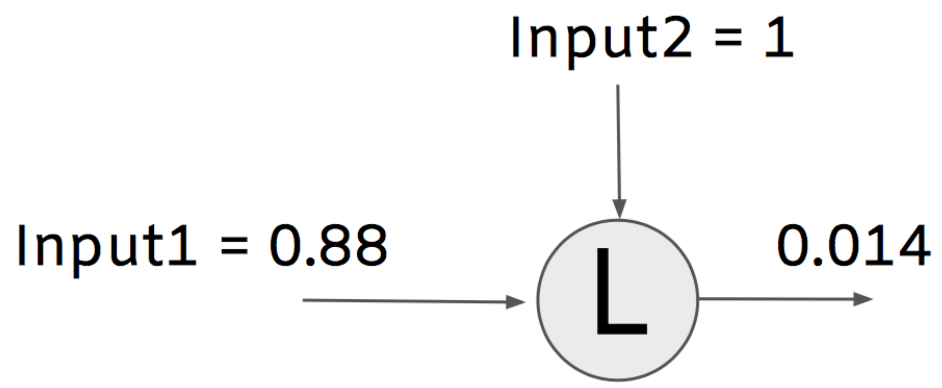
Chain rule!

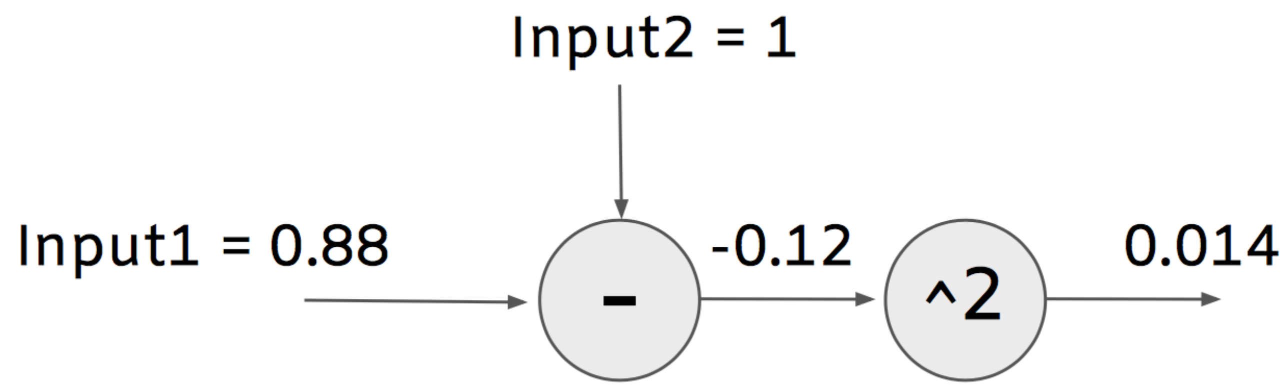
$$\frac{df(g(x))}{dx} = \frac{df(g(x))}{dg(x)} \frac{dg(x)}{dx}$$

MULTIPLY GRADIENTS



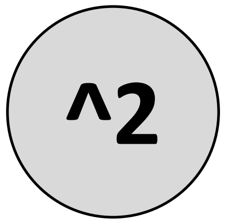
HOW IT WORKS





Loss = (Input1 - Input2)^2

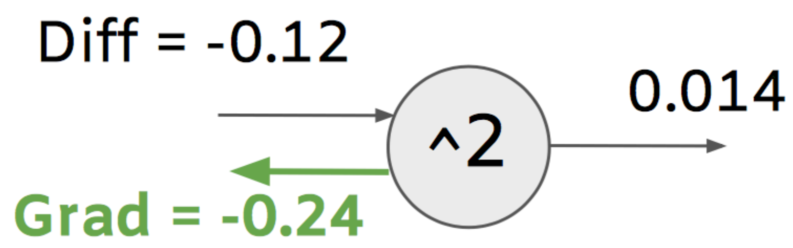
Derivative=2*Diff



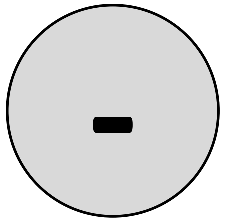
We can also write this as

Diff = Input1 - Input2

Loss = Diff^2

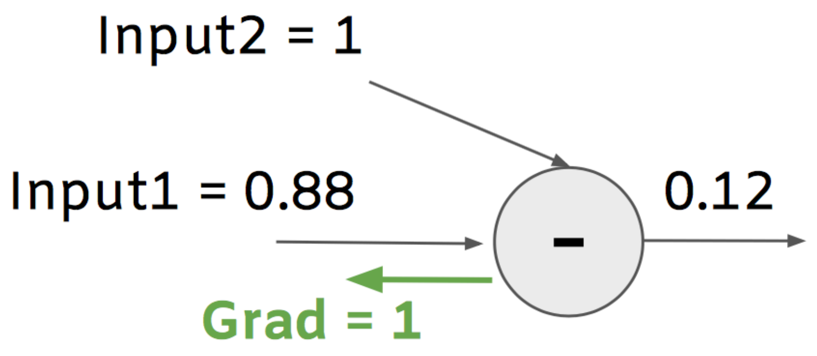


Derivative with Input1 = 1

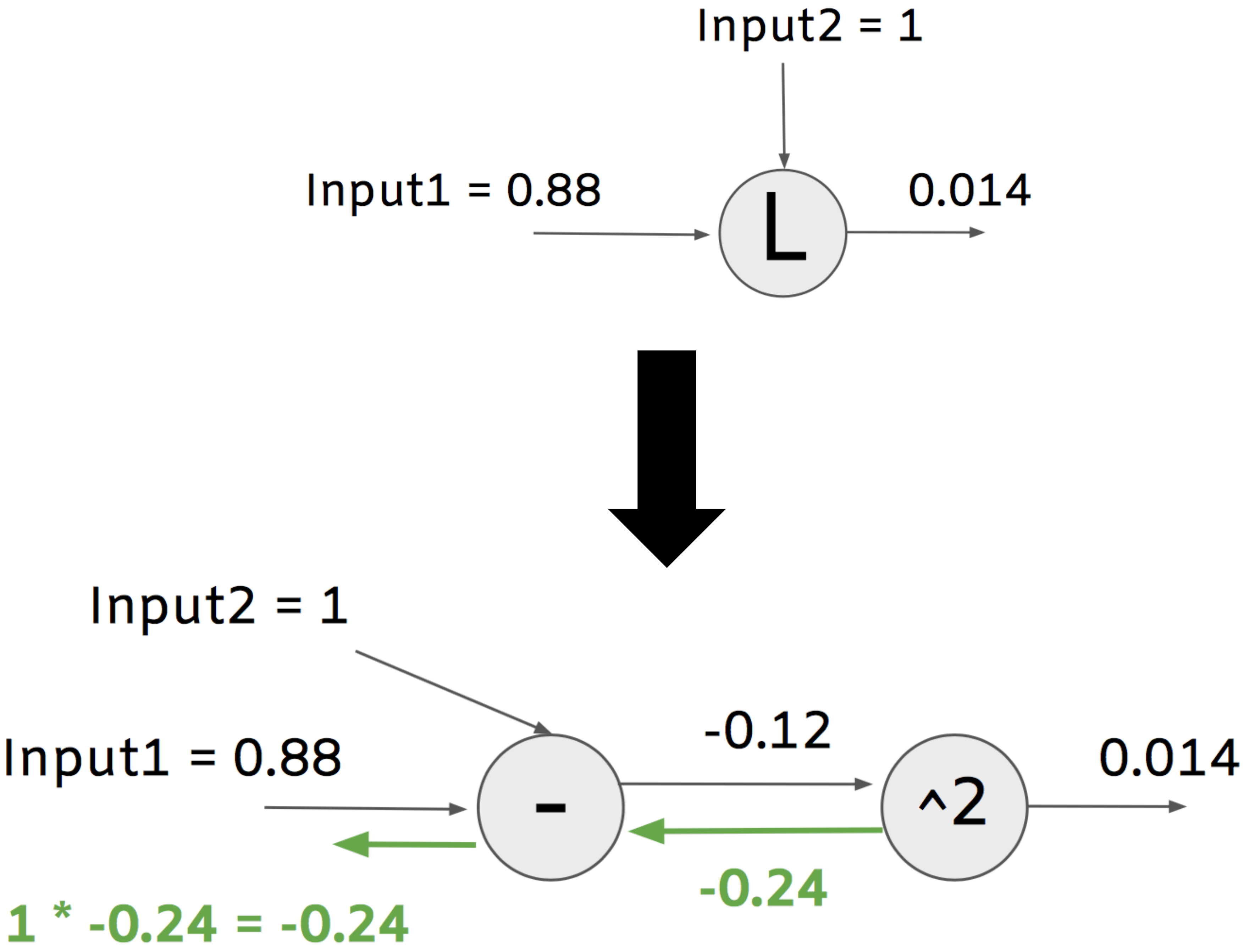


Output = Input1 - Input2

Derivative with Input1 = 1

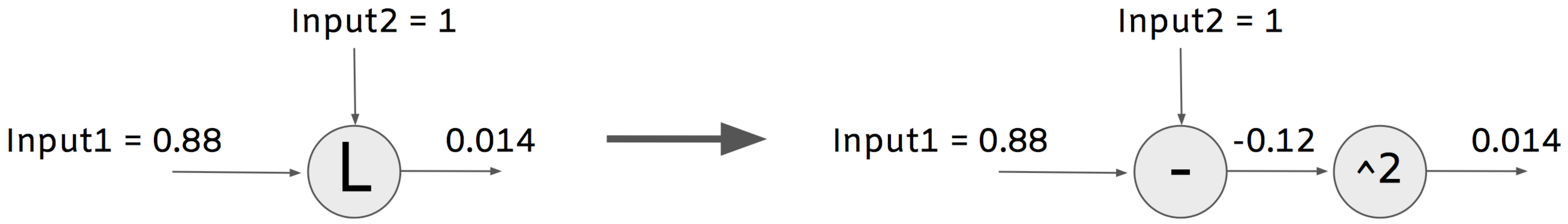


HOW IT WORKS



Input2 = 1
Input1 = 0.88
L
0.014
Input2 = 1
Input1 = 0.88
-
-0.12
^2
0.014
-0.24
1 * -0.24 = -0.24

HOW IT WORKS

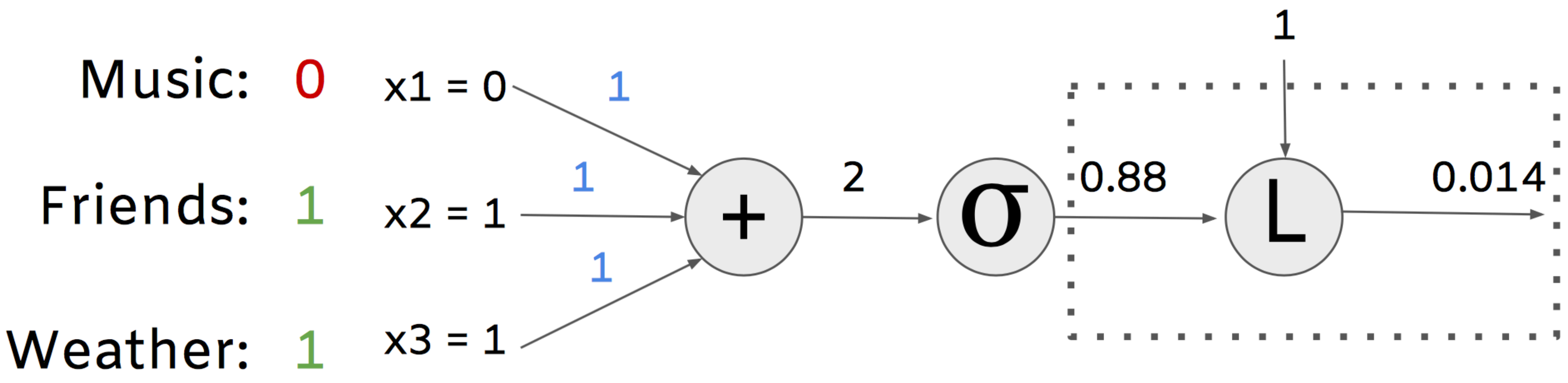


Loss = (Input1 - Input2)^2

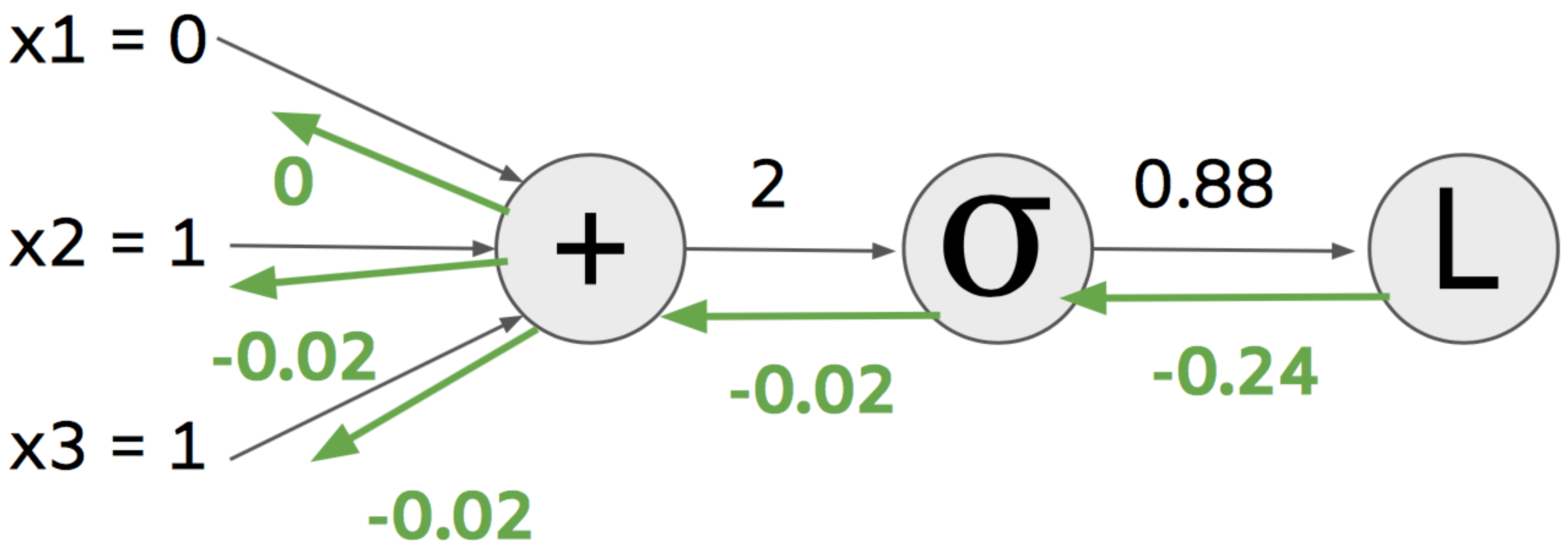
We can also write this as

Diff = Input1 - Input2

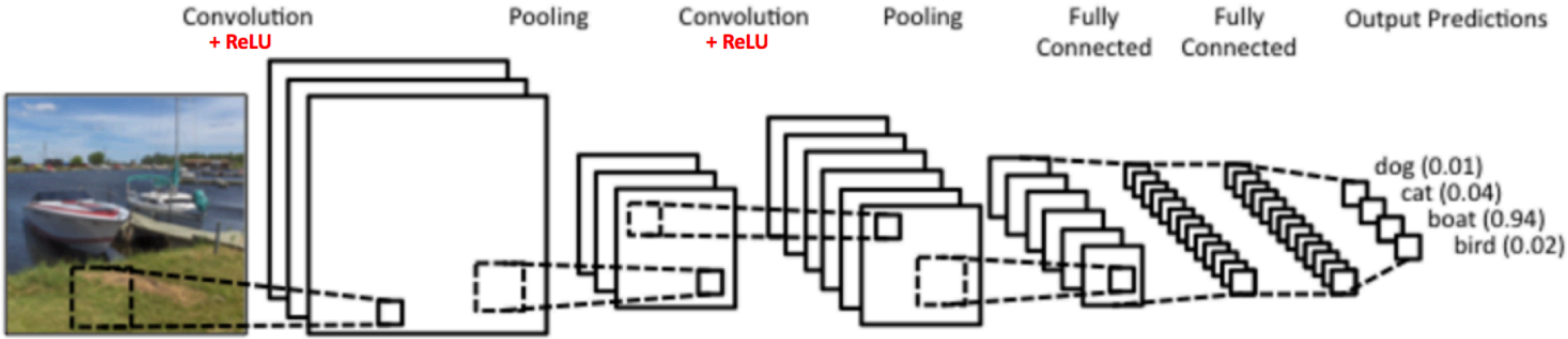
Loss = Diff^2



HOW IT WORKS



SIMPLE CNN



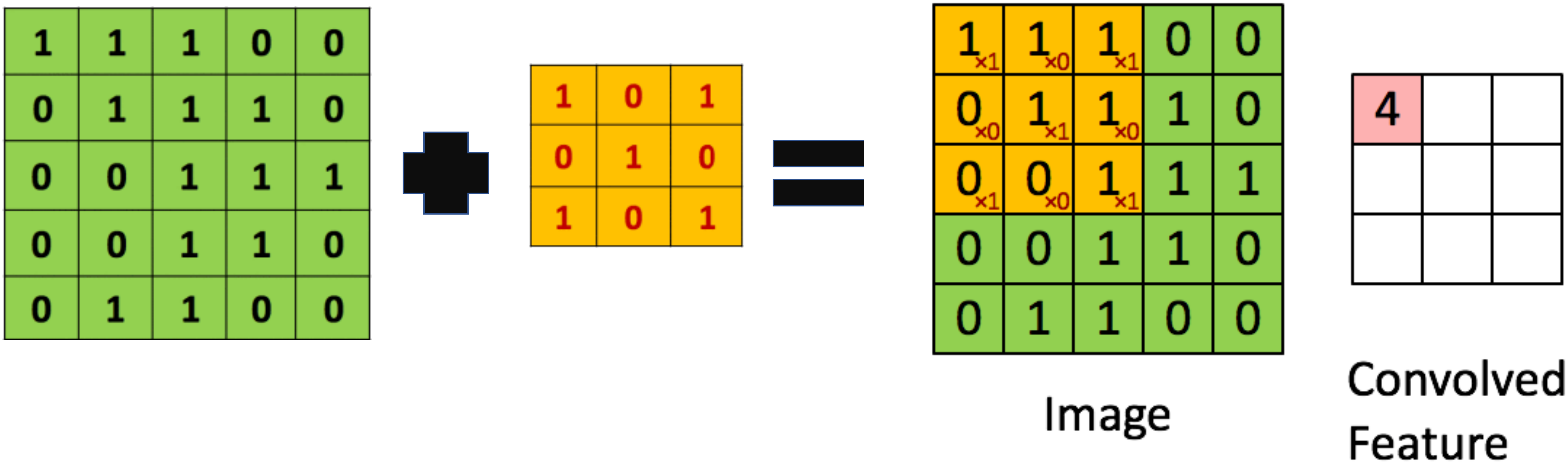
There are four main operations in the ConvNet shown above:

1. Convolution
2. Non Linearity (ReLU)
3. Pooling or Sub Sampling
4. Classification (Fully Connected Layer)

SIMPLE CNN - CONVOLUTION

The primary purpose of Convolution is to extract features from the input image.

The Convolution of the 5 x 5 image and the 3 x 3 matrix can be computed as shown as:



SIMPLE CNN - CONVOLUTION

The primary purpose of Convolution is to extract features from the input image.

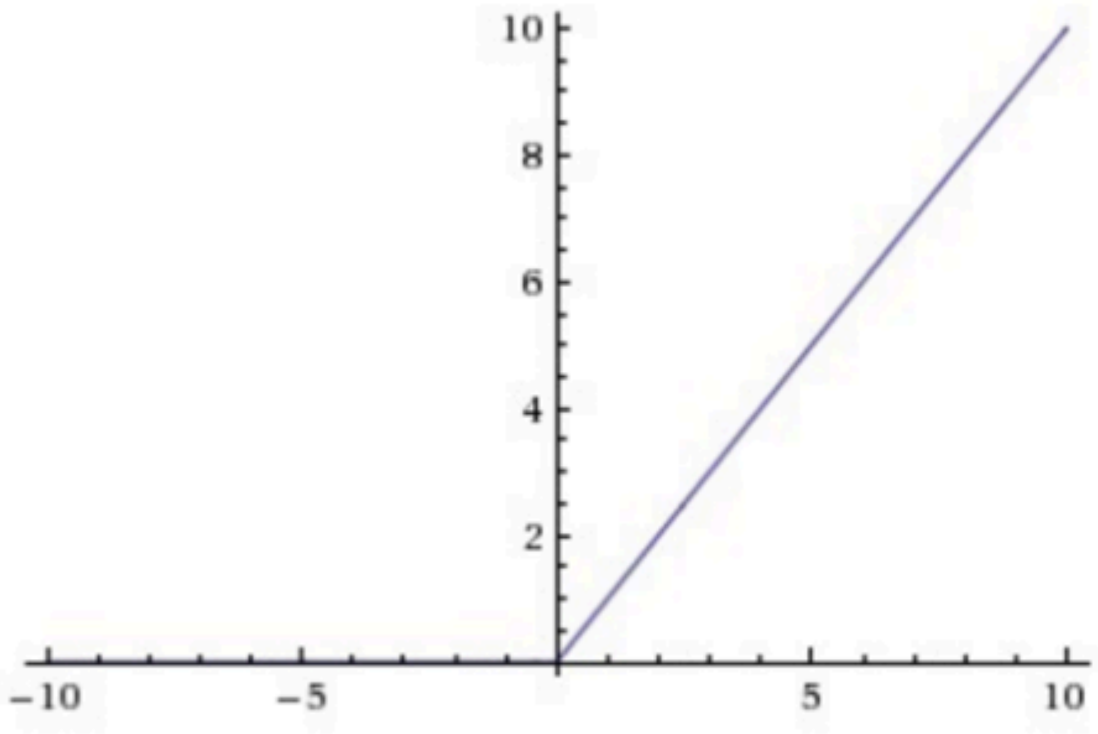
Operation	Filter	Convolved Image
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	

Operation	Filter	Convolved Image
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Gaussian blur (approximation)	$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

SIMPLE CNN - RELU

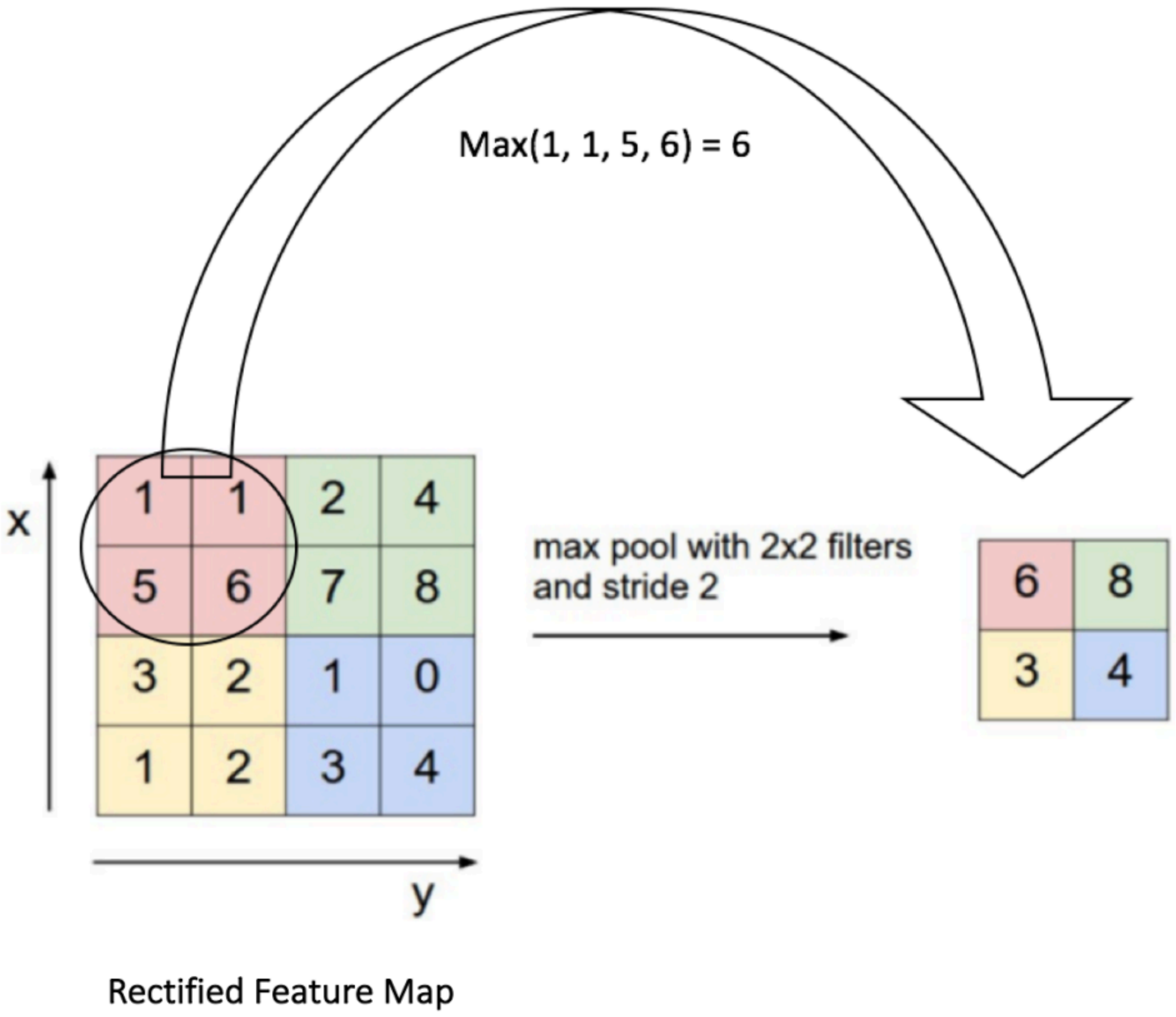
An additional operation called ReLU has been used after every Convolution operation. ReLU stands for Rectified Linear Unit and is a non-linear operation.

Output = Max(zero, Input)



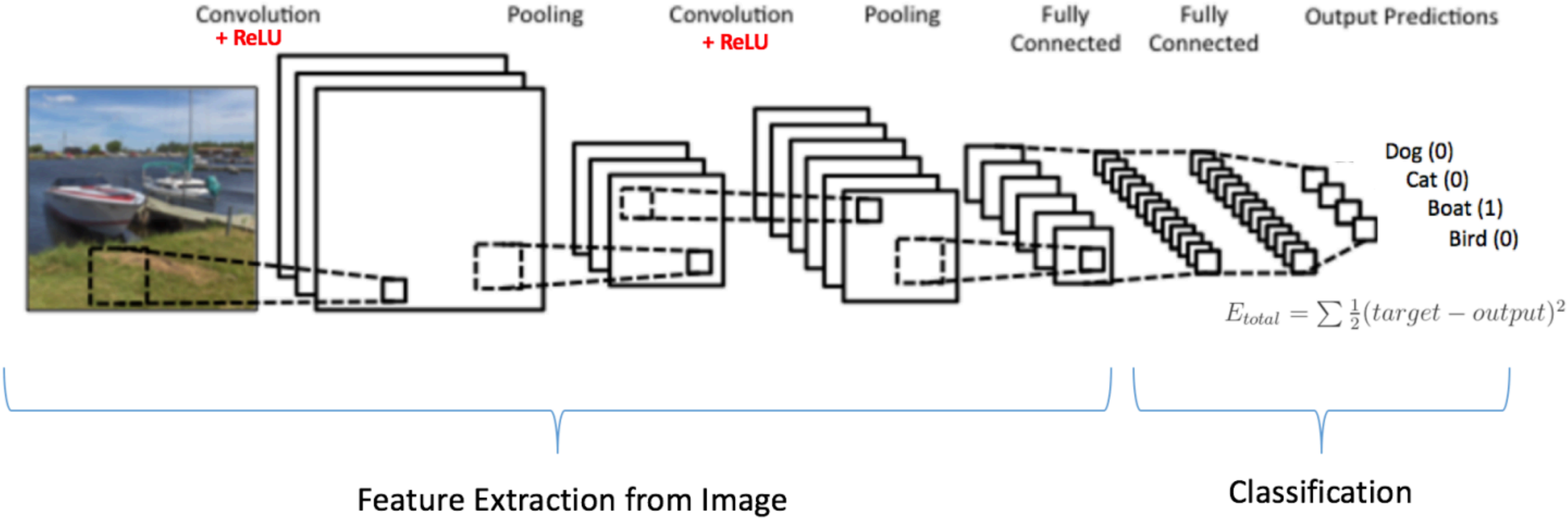
SIMPLE CNN - POOLING

Spatial Pooling (also called subsampling or downsampling) reduces the dimensionality of each feature map but retains the most important information. Spatial Pooling can be of different types: Max, Average, Sum etc.



SIMPLE CNN - CLASSIFICATION

The Fully Connected layer is a traditional Multi Layer Perceptron that uses a softmax activation function in the output layer (other classifiers like SVM can also be used). The term "Fully Connected" implies that every neuron in the previous layer is connected to every neuron on the next layer.



CNN Pros vs. Cons

Pros:

- Flexible
- Good for a variety of tasks
- Good for many types of data

Cons:

- Can require a lot of data
- Training may be slow
- Many parameters to tune
- Many layer types and activations
- Black Box model


Thank you!
Any Questions?