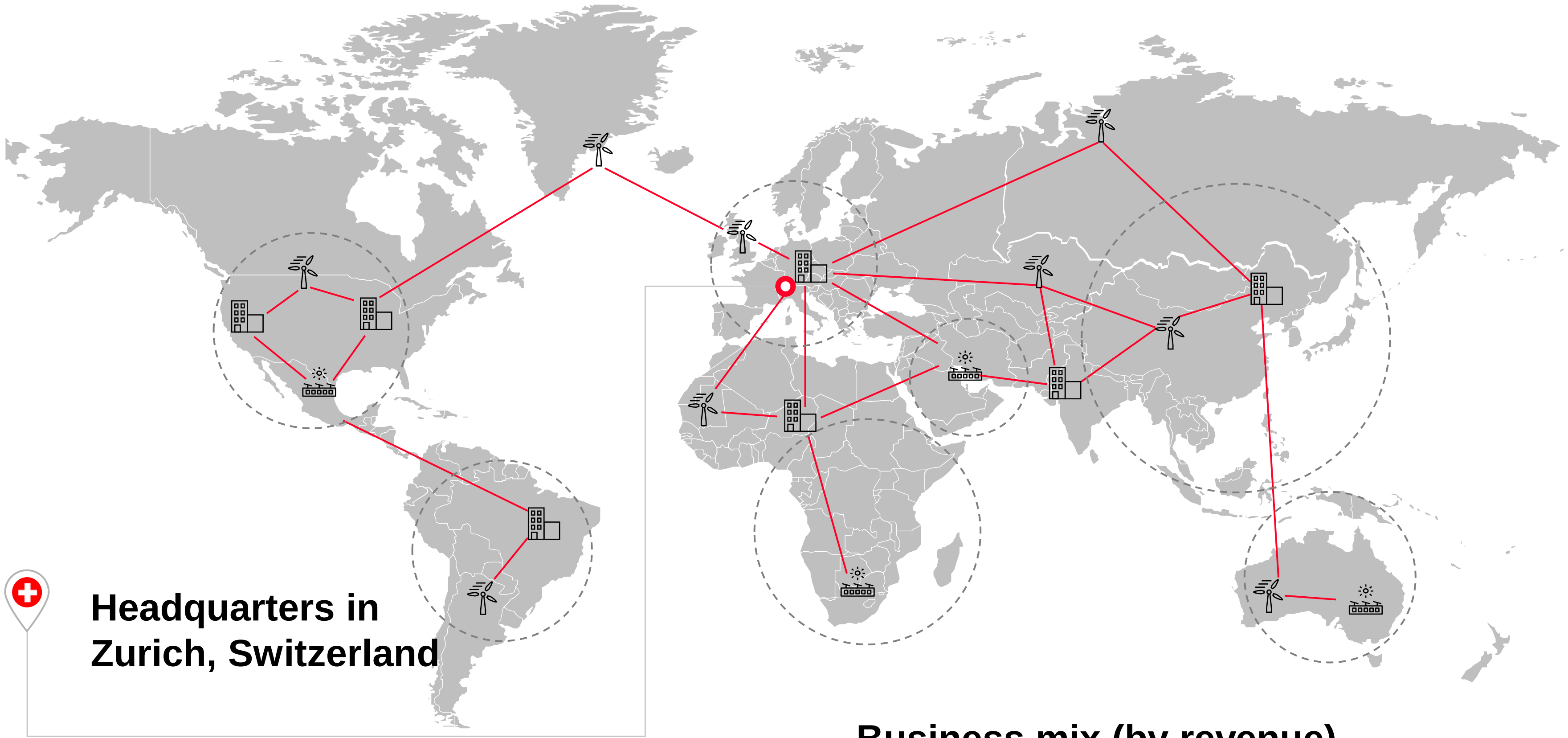
Reducing imbalance in time-series data

Industrial Problem-Solving Workshop 2023, Centre de Recherches Mathématiques

Jhelum Chakravorty

Hitachi Energy

About Hitachi Energy



38,000 employees

90+ countries with 200 offices

~250 years' heritage combined

5,500 sales employees & field engineers

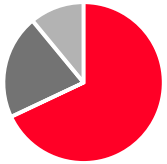
2,000 engineers & scientists in R&D

Four Business Units

Grid Automation	**High Voltage Products**
Grid Integration	**Transformers**

Business mix (by revenue)

Customers



- Transport & Infrastructure
- Industry
- Utilities

Offering

- Services
- Software & Automation
- Systems
- Products

Geographies

- Asia, Middle East & Africa
- Americas
- Europe

Hitachi Energy

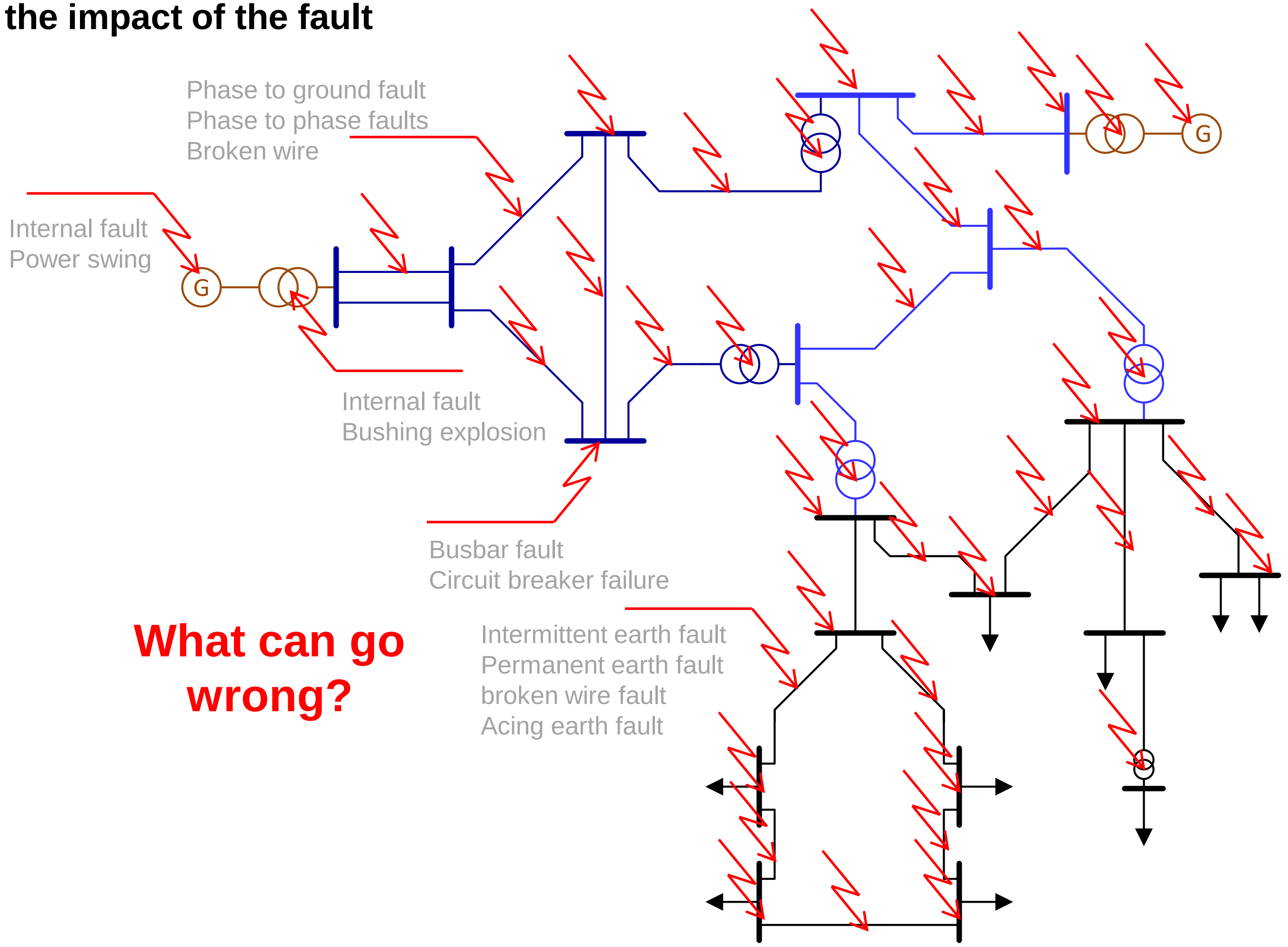
Case study: Power system protection

HITACHI
Inspire the Next

Isolate the faulted section to protect the other parts from the impact of the fault

Goal and requirements

- **Speed** – isolate fault as soon as possible in order to limit equipment damage
- **Selectivity** – isolate only faulted section of the power system
- **Sensitivity / Security** – isolate every fault
- **Reliability / Dependability** – only act when appropriate (be insensitive to typical operation events OR faults in close vicinity but that should be cleared by other breakers)

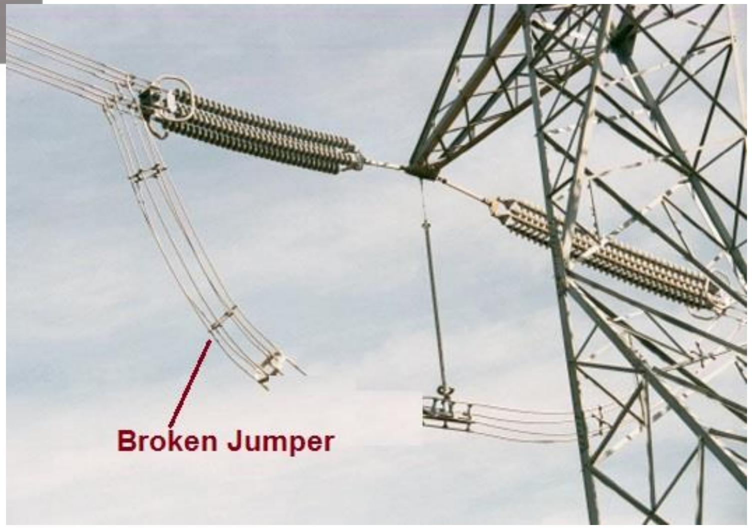


Power system protection is the last line of defense when equipment is physically endangered due to a fault event or operating conditions being out of acceptable limits

Hitachi Energy

Case study: Power system protection

HITACHI
Inspire the Next



Goal and requirements

- **Speed** – isolate fault as soon as possible in order to limit equipment damage
- **Selectivity** – isolate only faulted section of the power system
- **Sensitivity / Security** – isolate every fault
- **Reliability / Dependability** – only act when appropriate (be insensitive to typical operation events OR faults in close vicinity but that should be cleared by other breakers)

Power system protection is the last line of defense when equipment is physically endangered due to a fault event or operating conditions being out of acceptable limits

Hitachi Energy

Power system protection – Improve tripping of relay in overhead transmission line

HITACHI
Inspire the Next

ML solution case study

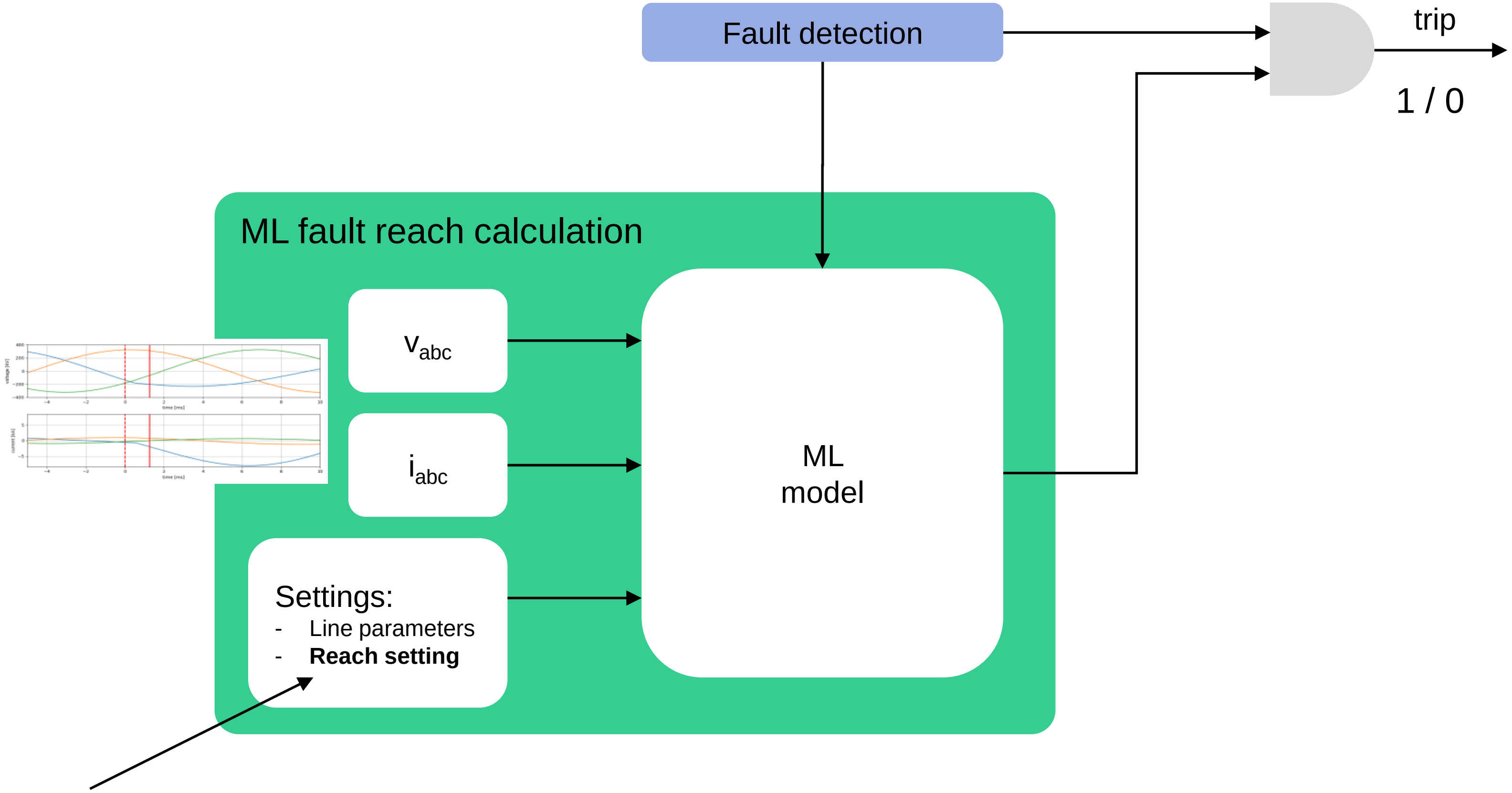
- Fault detection within a few samples after occurrence of fault
- Reach setting (protected zone) information
- Binary decision of trip (1) or restrain (0)

Requirements:

- Operate under 4.8kHz real time system
- Decision made based only on past samples
- Decision made for each time-sample

Priorities (order or preference):

- High dependability (no false positive)
- High speed (~ms)
- High security (no false negatives)



Data imbalance depends heavily on this

ML model requires to make a binary decision every sample if circuit breaker should be tripped

Hitachi Energy

ML for protection - Why is it challenging?

HITACHI
Inspire the Next

Example case

ML model makes a decision if a transmission line should be tripped

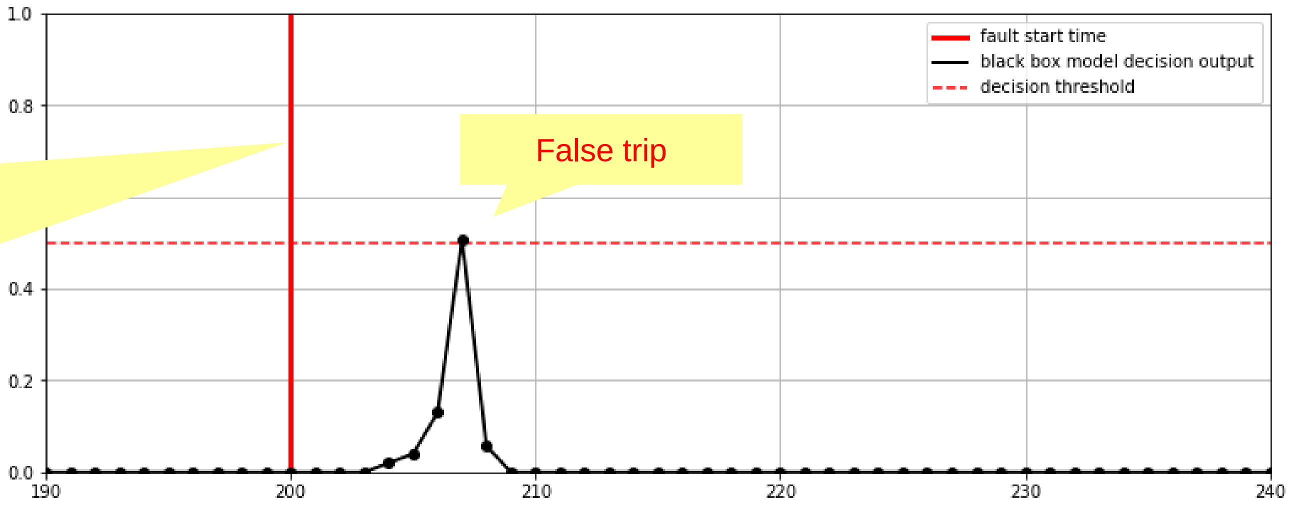
Machine learning: sequential decision using only past samples at each step to decide if circuit breaker should trip a fault within protected zone

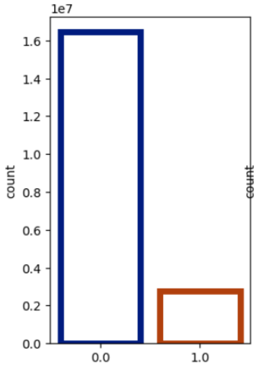
- maintain causality; no peaking in the future samples
- 0.x % error (MSE)
- 99.x % accuracy
- Degree of imbalance depends on reach setting
- False positive more detrimental than false negative

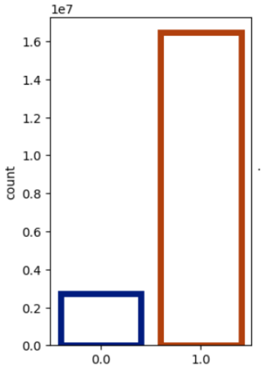
Power system Protection:

- Speed = 1.45 ms
- Mal-operation

Fault outside protected zone (e.g., in a neighboring line), should not trip this relay







Left: reach setting range (0.0 – 0.33)

Right: reach setting range (0.67 – 1.00)

Trip: red
Restrain: Blue

Power system protection require to solve causal and temporal problem of high-stake decision making

Hitachi Energy

ML for protection – simulation setup and dataset

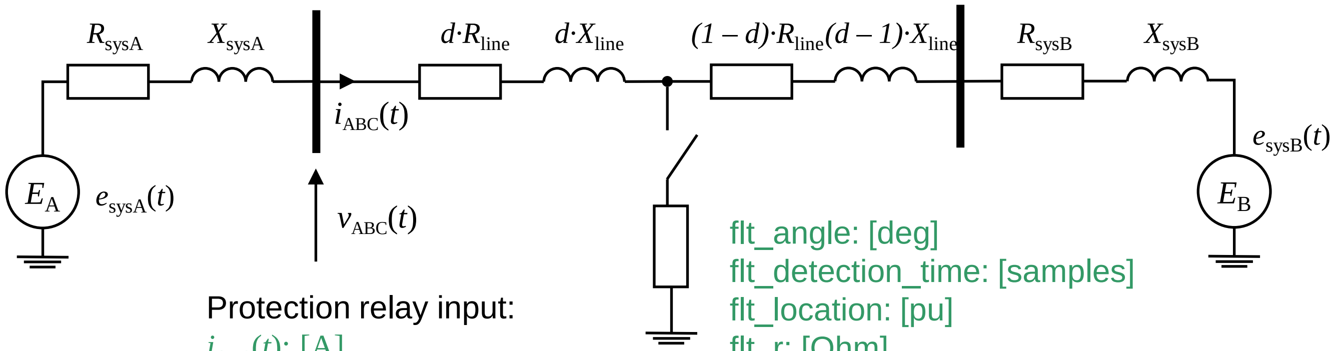
50000 cases, 27 features, 576 time-samples

Simulation setup

line_len: [km]
line_r0: [Ohm/km]
line_r1: [Ohm/km]
line_x0: [Ohm/km]
line_x1: [Ohm/km]
line_c0: [F/km]
line_c1: [F/km]

sys_b_r0 [Ohm]
sys_b_r1 [Ohm]
sys_b_sir: [pu]
sys_b_x0: [Ohm]
sys_b_x1: [Ohm]
sys_b_z0z1_ratio: [pu]
load_angle: [deg]

sys_a_V_nom: [V]
sys_a_r0 [Ohm]
sys_a_r1 [Ohm]
sys_a_sir: [pu]
sys_a_x0: [Ohm]
sys_a_x1: [Ohm]
sys_a_z0z1_ratio: [pu]



flt_angle: [deg]
flt_detection_time: [samples]
flt_location: [pu]
flt_r: [Ohm]

Protection relay input:
$i_{ABC}(t)$: [A]
$v_{ABC}(t)$: [V]
line_r0: [Ohm]
line_r1: [Ohm]
line_x0: [Ohm]
line_x1: [Ohm]
reach_setting: [pu] (default 0.7)

Operate if flt_location < reach_setting as fast as possible after fault detection

Going beyond resampling for reducing data imbalance

- State-of-the-art resampling methods provide ways to up / down sample minority / majority class representation
  - Usually, they are suitable for *independently and identically distributed* data samples
  - *hyper-parameters* (e.g., stopping criteria) that need to be relearnt for each use case
  - They are *costly and slow* for extremely skewed imbalance

- Time series data with power system protection has more acute degree of imbalance in the time dimension
- **Causality** needs to be respected; sequential decision at time steps
- Decision has to be **fast and accurate**

- **We are looking for a methodical approach to address efficient reduction of imbalance in time series data to successfully render fast and accurate decision**

Thank you

Hitachi Energy

HITACHI
Inspire the Next