A federated learning approach to mixed fault diagnosis in rotating machinery

Manan Mehta[1], Siyuan Chen[1], Haichuan Tang[2], Chenhui Shao[1]

[1]Mechanical Science and Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
[2]CRRC Academy, Beijing 100161, China

North American Manufacturing Research Conference (NAMRC) 51

June 14th 2023

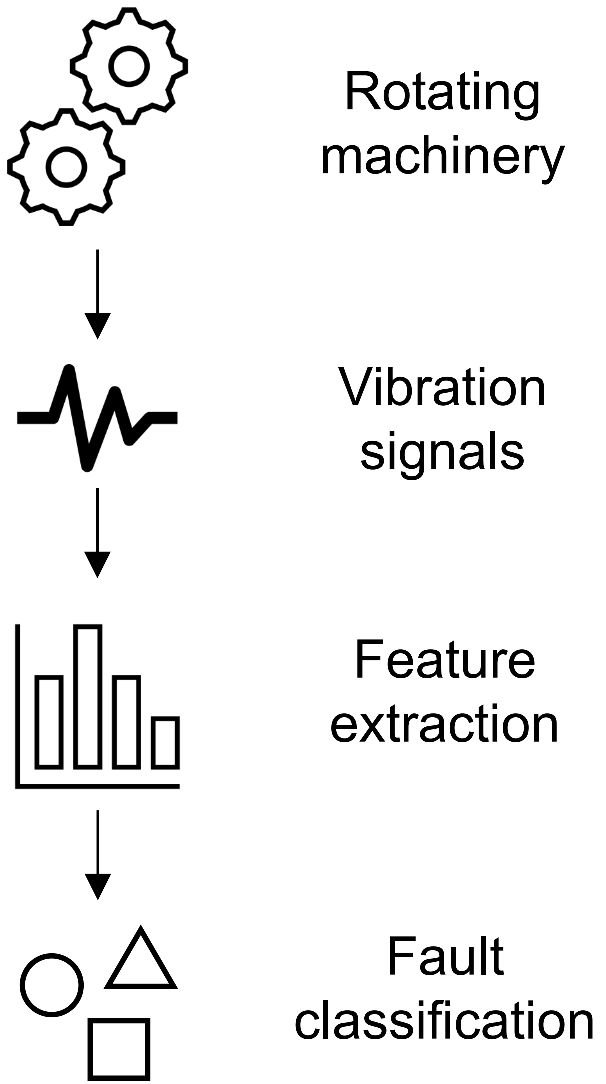
Rutgers University, New Brunswick, NJ, USA

Introduction

- Ensuring optimal operating conditions for rotating machinery is essential in industrial applications

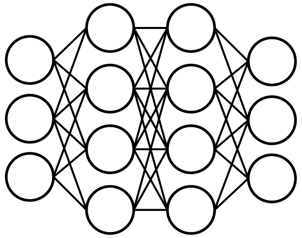
- Fault diagnosis methods can be:
  - Analytical
  - Knowledge/physics-driven
  - Data-driven

- Data-driven deep learning (DL) methods for fault diagnosis from vibration signals are most popular



Introduction

DL-based fault diagnosis in literature:



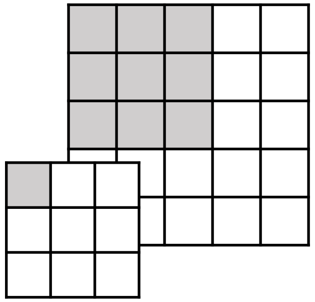
Multi-layer perceptron

[Chen and Mo, 2004]

[Rafiee et al., 2007]

[Bin et al., 2012]

[Chandra and Sekhar, 2016]



Convolutional neural network

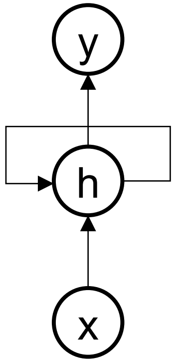
[Janssens et al., 2016]

[Xia at al., 2017]

[Guo et al., 2018]

[Chen et al., 2020]

[Li et al., 2020]



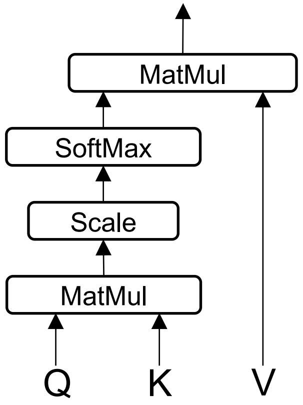
Recurrent neural network and LSTM

[Yuan et al., 2016]

[Yang et al., 2018]

[Jalayer et al., 2021]

[Zhang et al., 2021]



Attention and transformer

[Pei et al., 2021]

[Zhao et al., 2021]

[Jin et al., 2022]

[Shao et al., 2023]

Introduction

- ❑ Performance of data-driven DL algorithms depends on the quality and quantity of training data
- ❑ Collecting, labeling, and storing sensor data is resource-intensive for individual factories
- ❑ Similar data at other factories cannot be pooled due to its sensitive nature
- ❑ Two main bottlenecks for DL-based fault diagnosis:

Data Availability

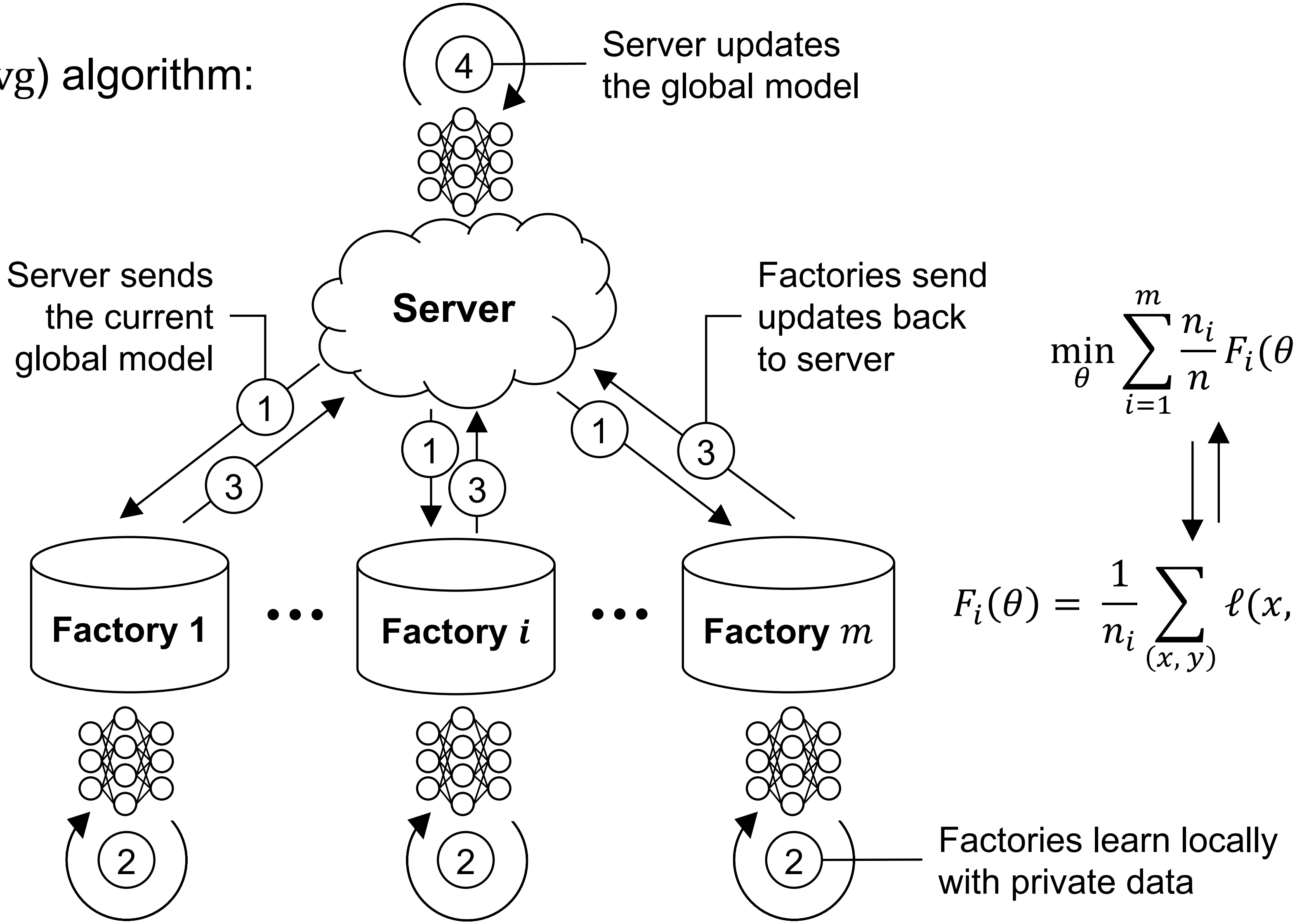
Data Privacy

Federated learning (FL) allows multiple manufacturers to build a collaborative DL model while keeping their training data private

Federated learning: FedAvg algorithm

The Federated Averaging (FedAvg) algorithm:

- ❑ Privacy advantage over centralized learning
- ❑ Ability to handle non-IID data
- ❑ Ability to learn over unbalanced datasets

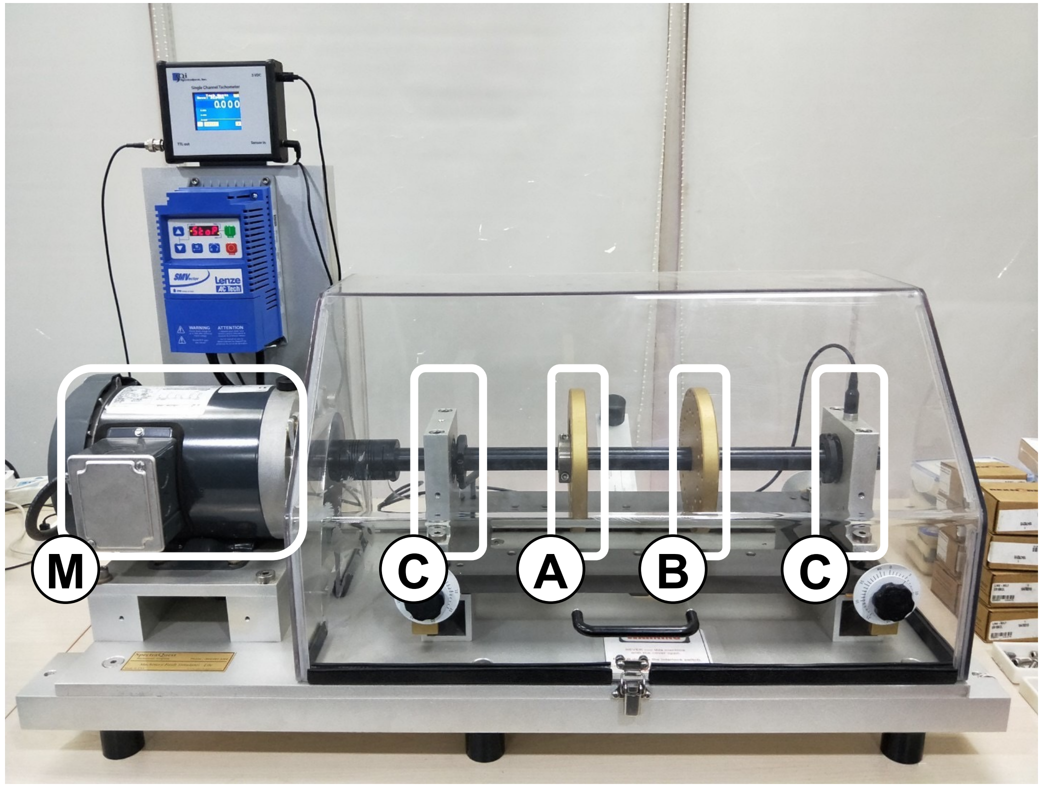


$$\min_{\theta} \sum_{i=1}^{m} \frac{n_i}{n} F_i(\theta)$$

$$F_i(\theta) = \frac{1}{n_i} \sum_{(x,y)} \ell(x, y|\theta)$$

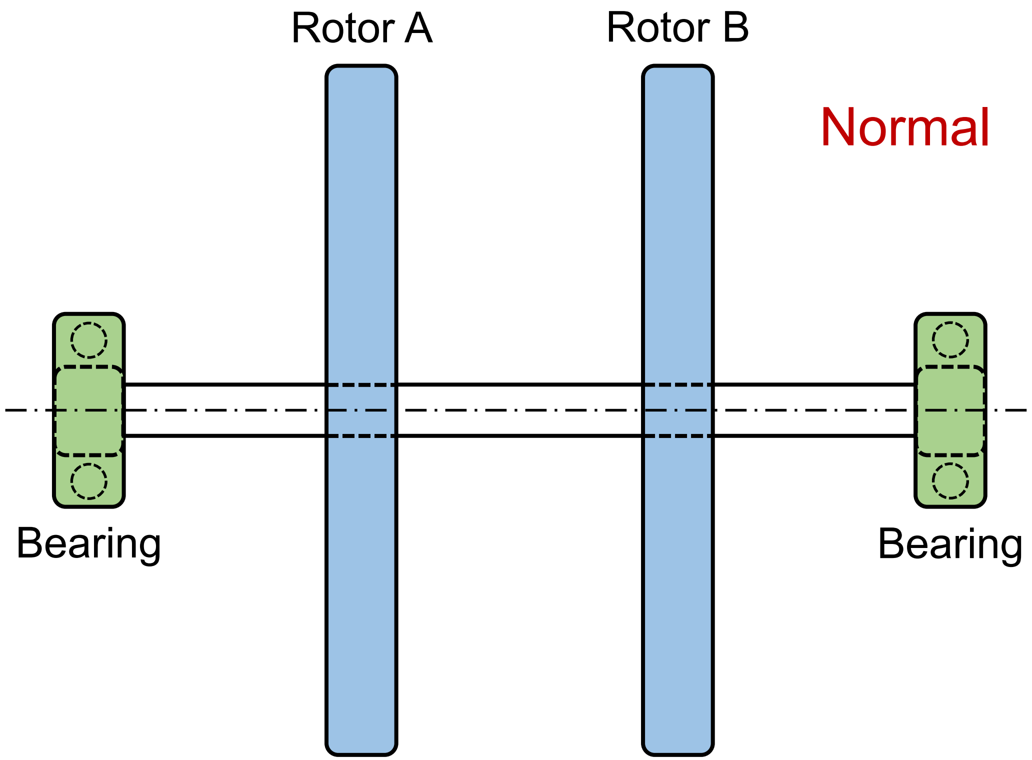
Data collection and preprocessing

- A specialized machinery fault simulator (MFS) used to collect mixed fault signals
- MFS consists of
  - Motor
  - Two bearings
  - Two rotors
  - Tachometer (rotating speed)
  - Accelerometer (lateral vibrations)
- A combination of six rotor and eight bearing conditions result in 48 total machine health states
- A total of 82 hours worth of data collected



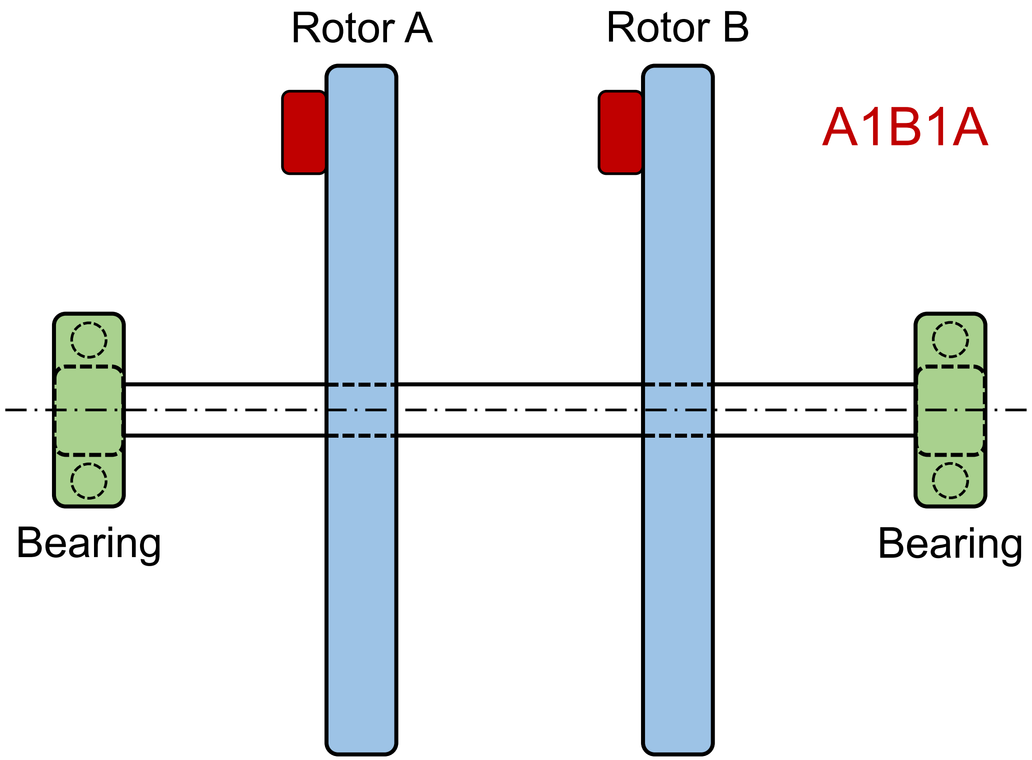
(M) Motor (A) (B) Rotors (C) Bearings

Data collection and preprocessing



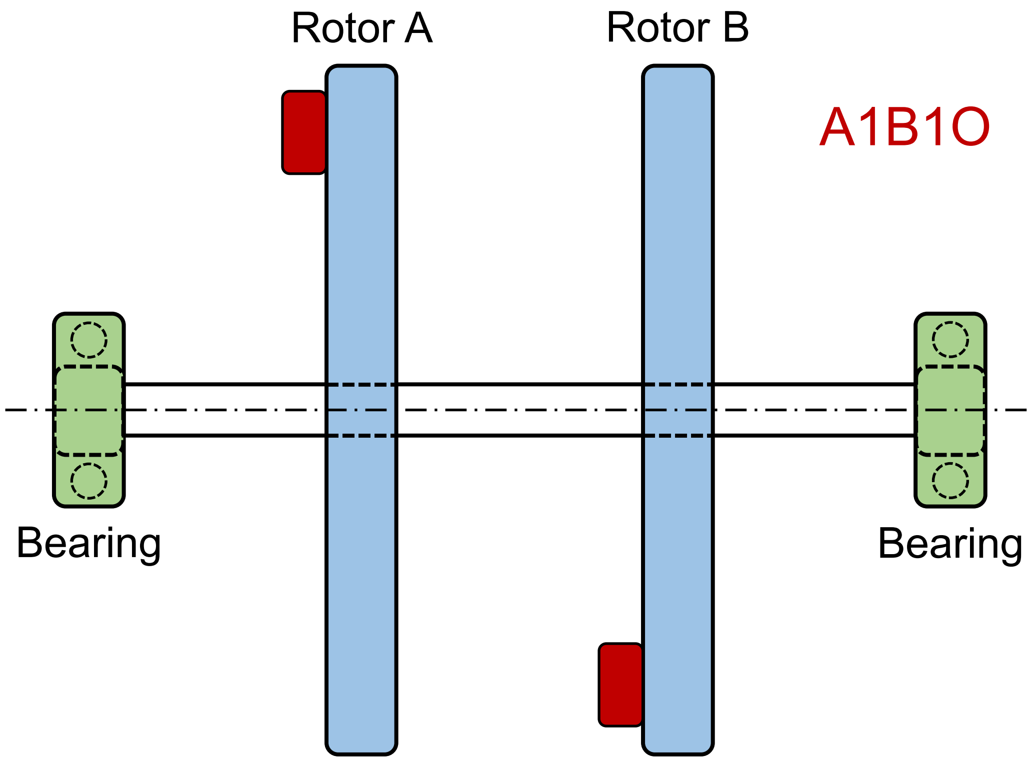
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



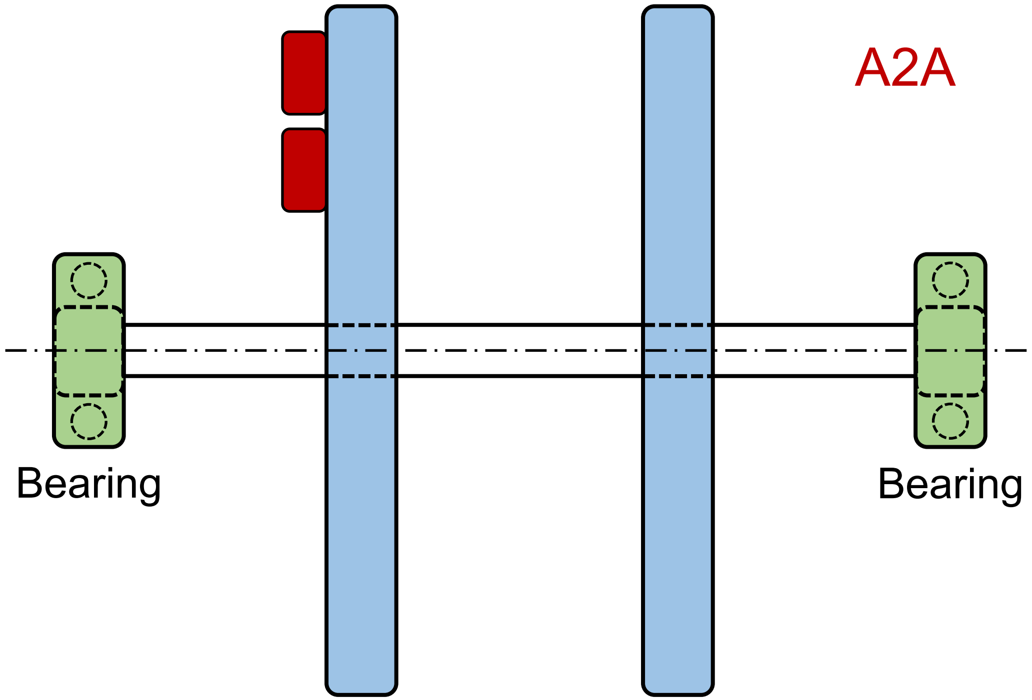
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



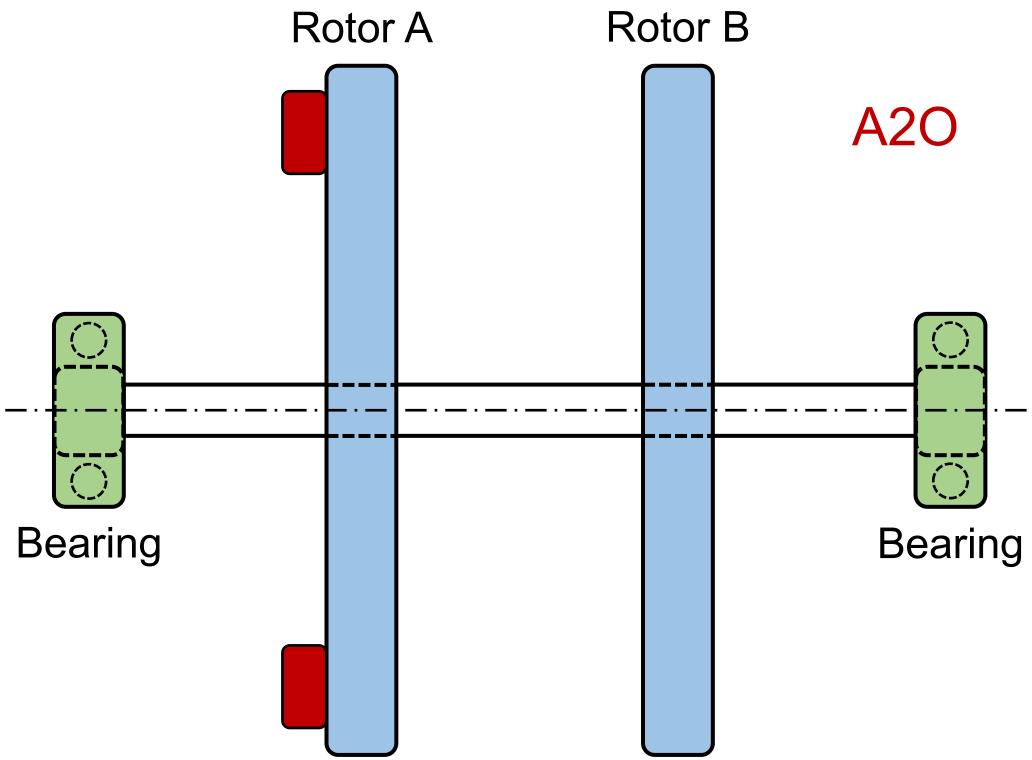
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



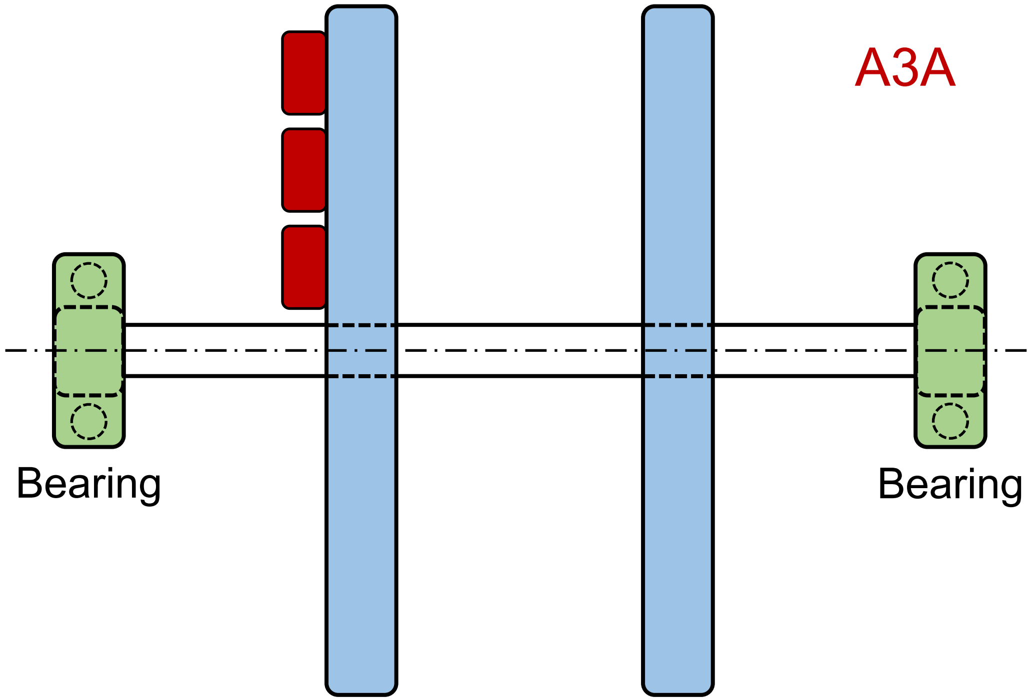
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



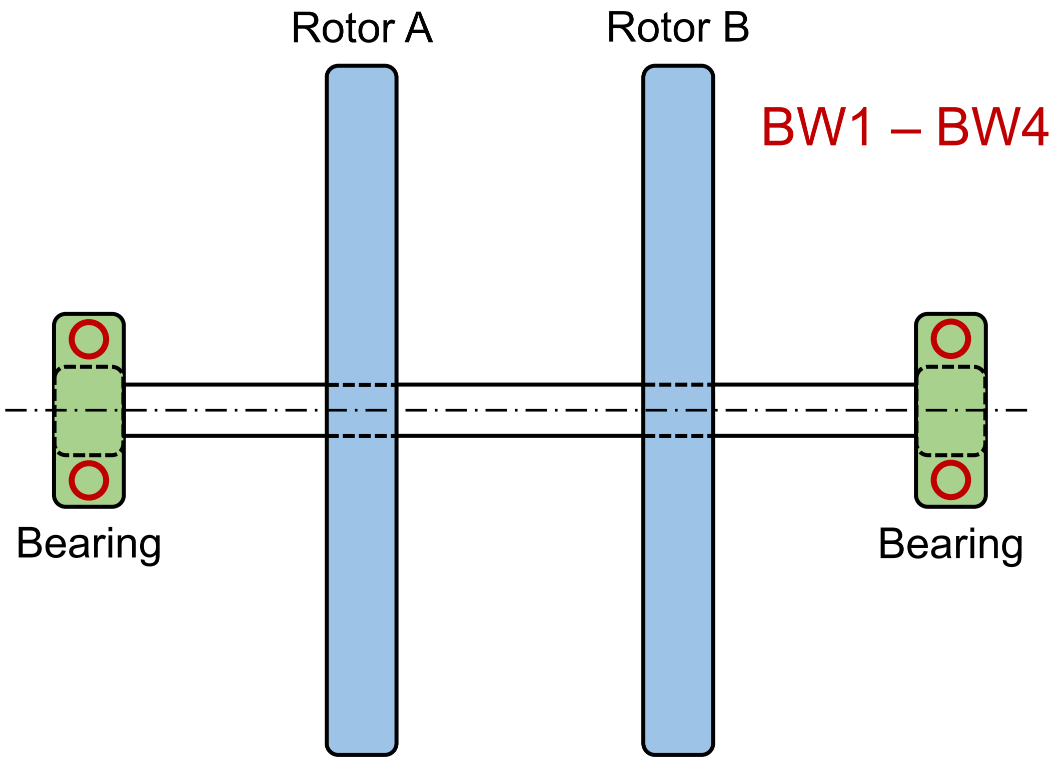
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



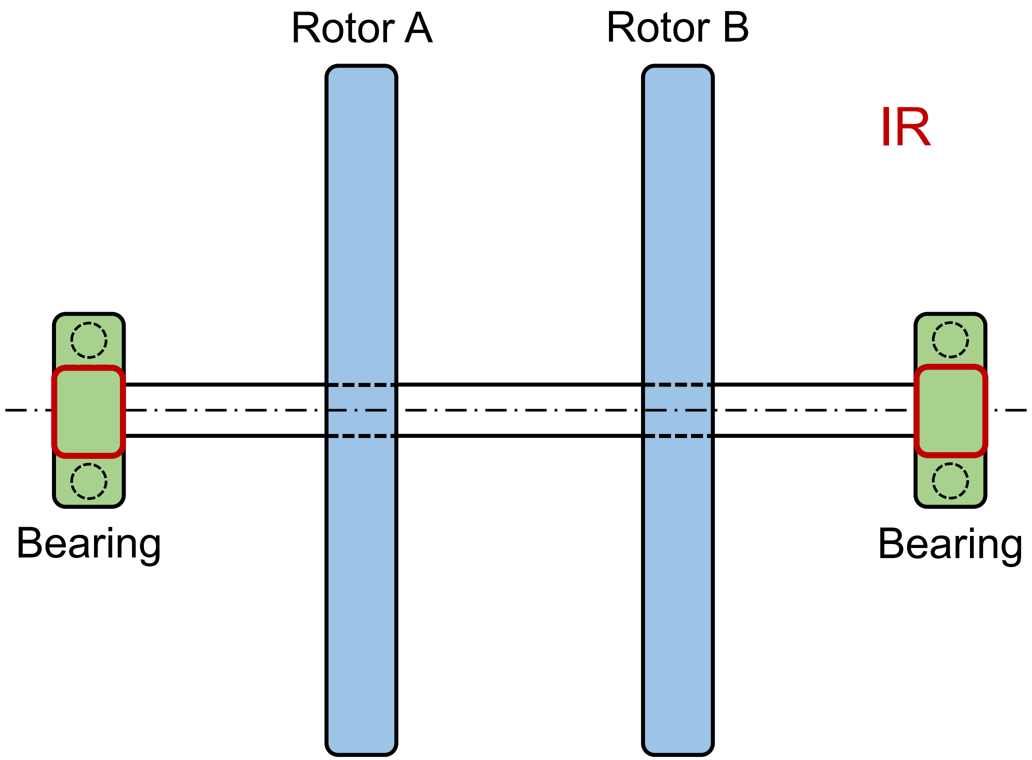
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



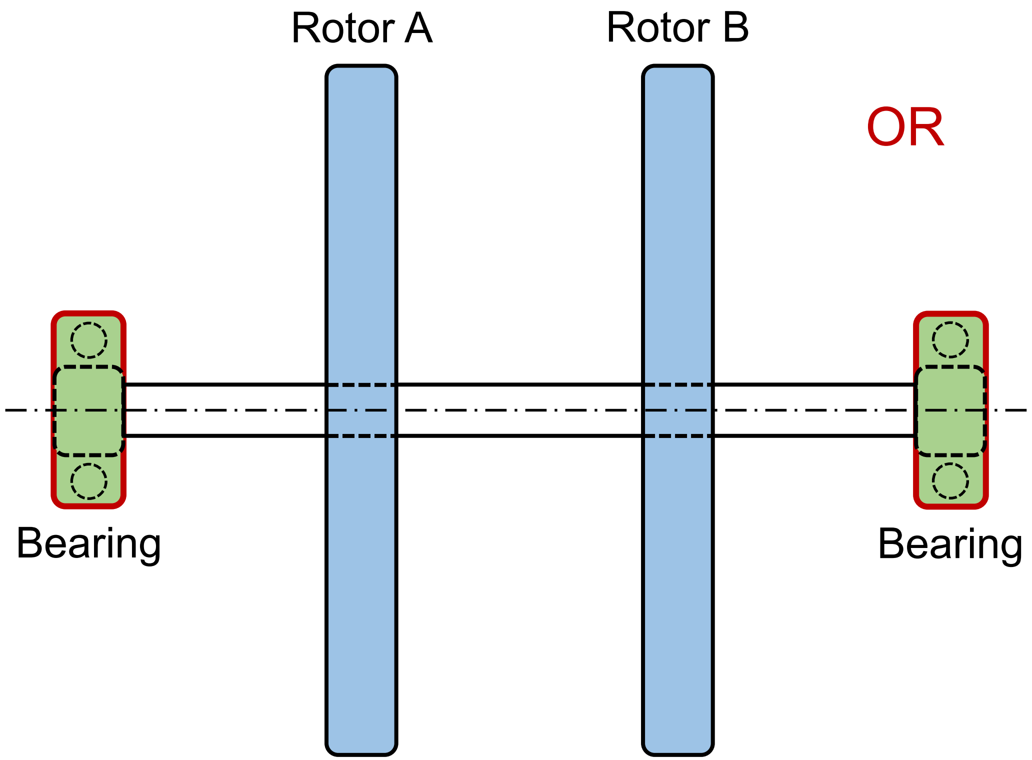
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



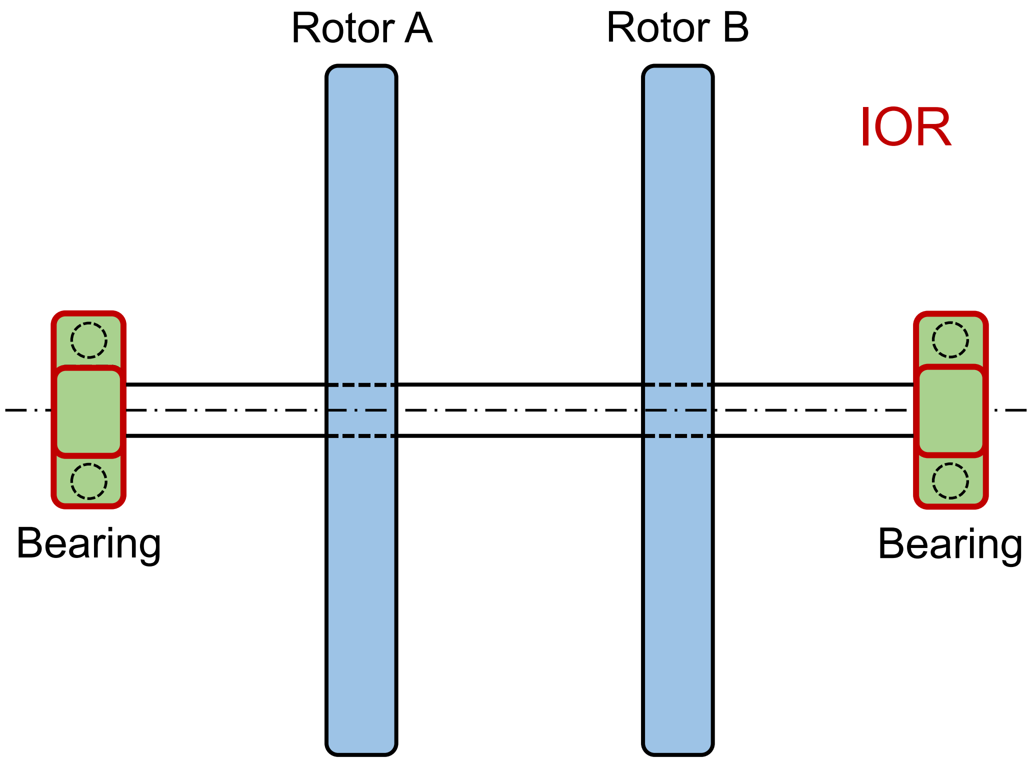
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



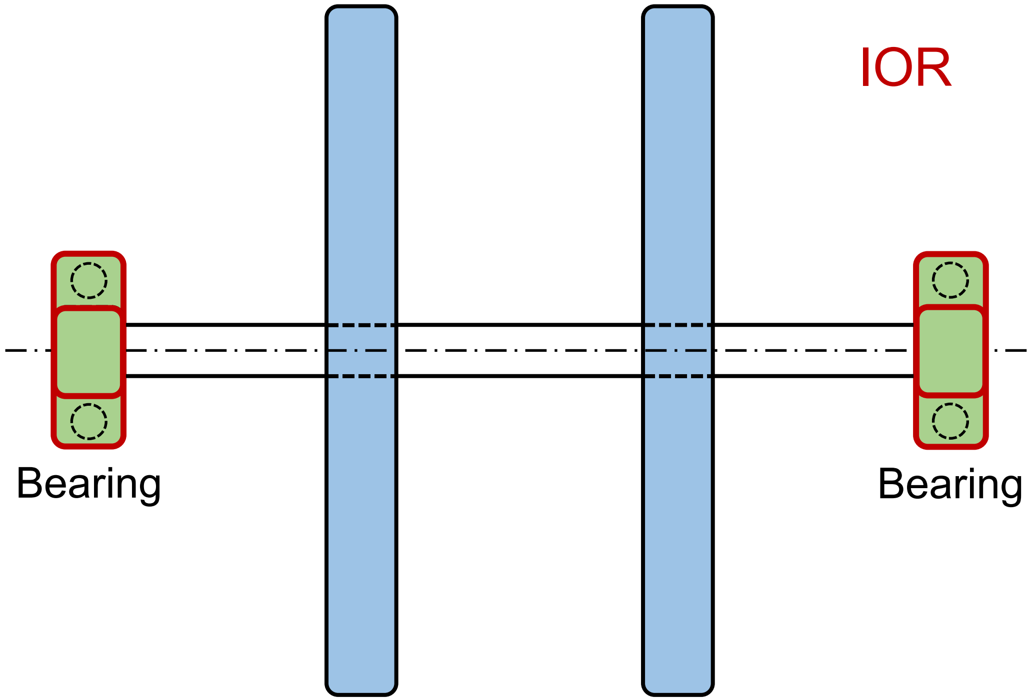
Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

Data collection and preprocessing



Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

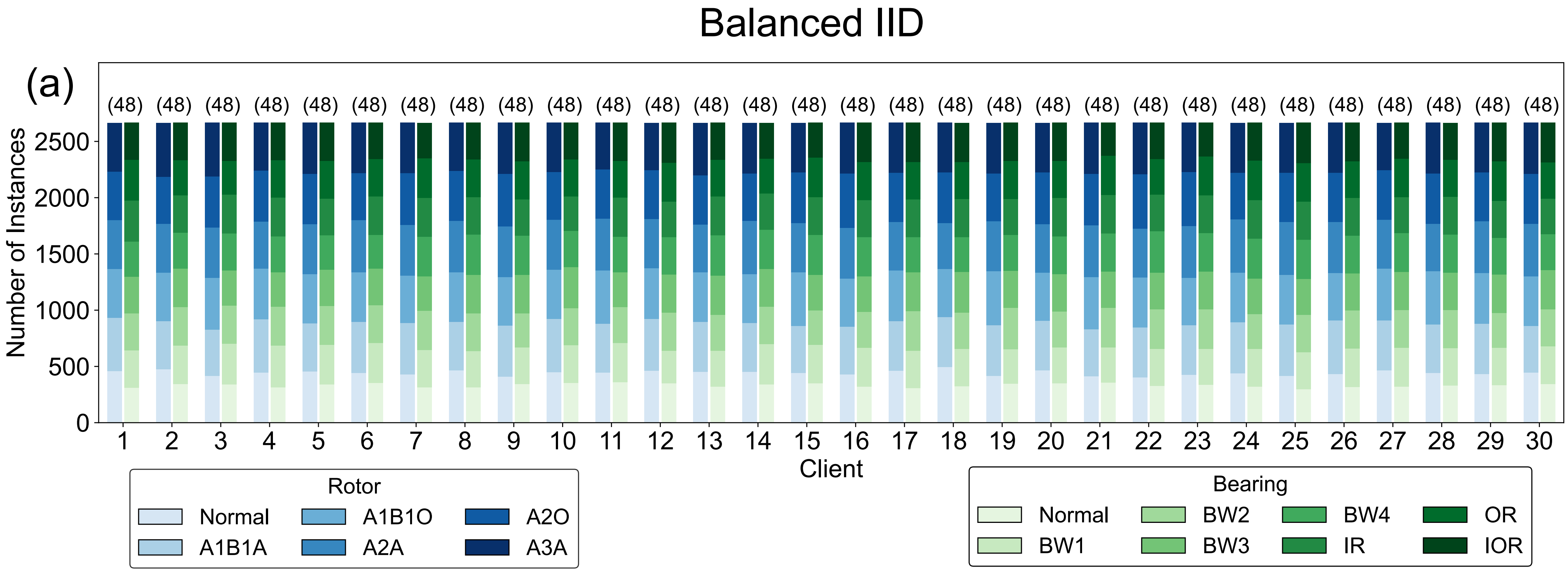
Data collection and preprocessing



Class	Rotor	1	2	3	4	5	6
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
1	Normal	1	2	3	4	5	6
2	BW1	7	8	9	10	11	12
3	BW2	13	14	15	16	17	18
4	BW3	19	20	21	22	23	24
5	BW4	25	26	27	28	29	30
6	IR	31	32	33	34	35	36
7	OR	37	38	39	40	41	42
8	IOR	43	44	45	46	47	48

- ❑ Signals collected at 720, 840, 960, 1080, 1200 RPM, then interpolated to 600 RPM
- ❑ 1920 signals for each class x 48 classes = 92,160 signals in total

Case study: Data distributions

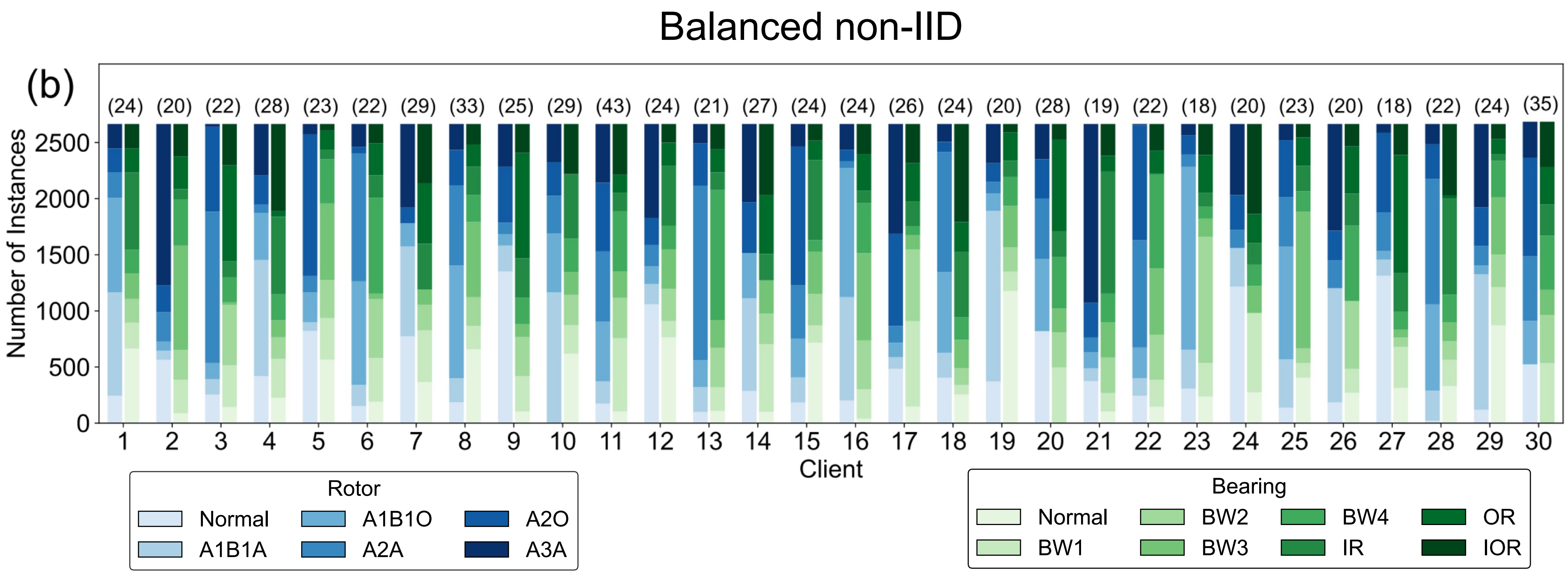


Balanced IID

All clients have all mixed fault labels

All clients have equal number of samples

Case study: Data distributions



Balanced IID

All clients have all mixed fault labels

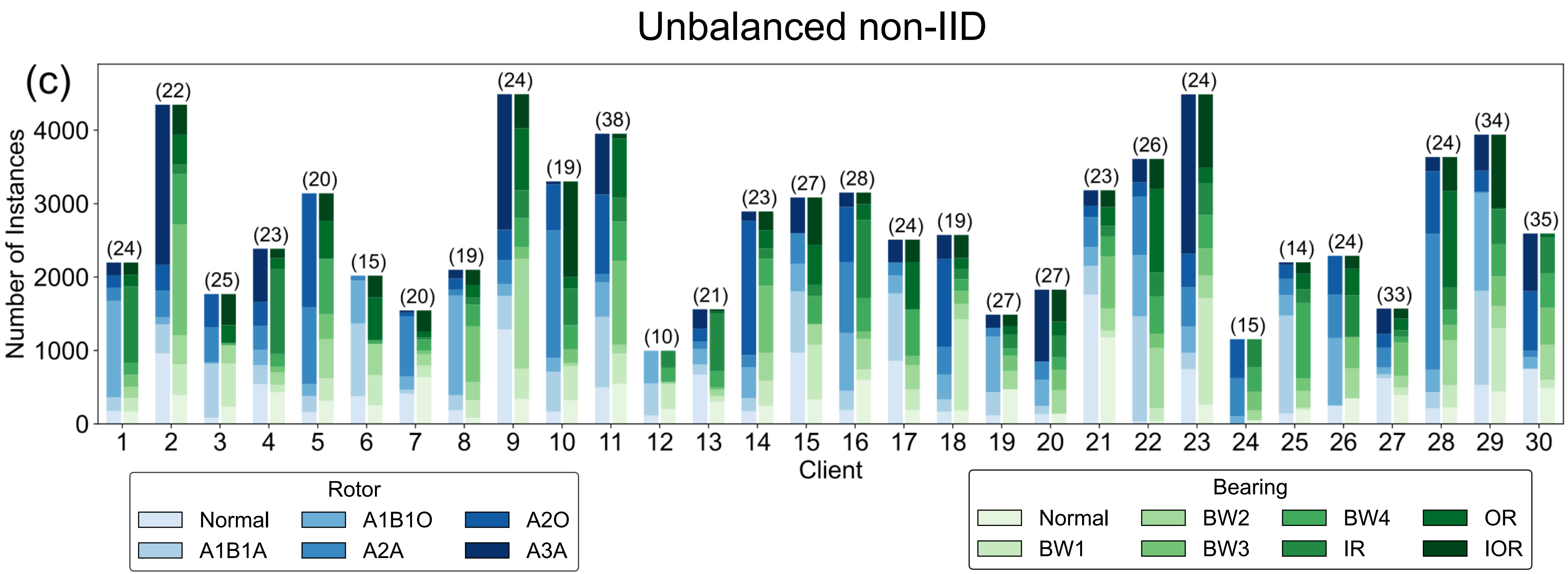
All clients have equal number of samples

Balanced non-IID

All clients do not have all mixed fault labels

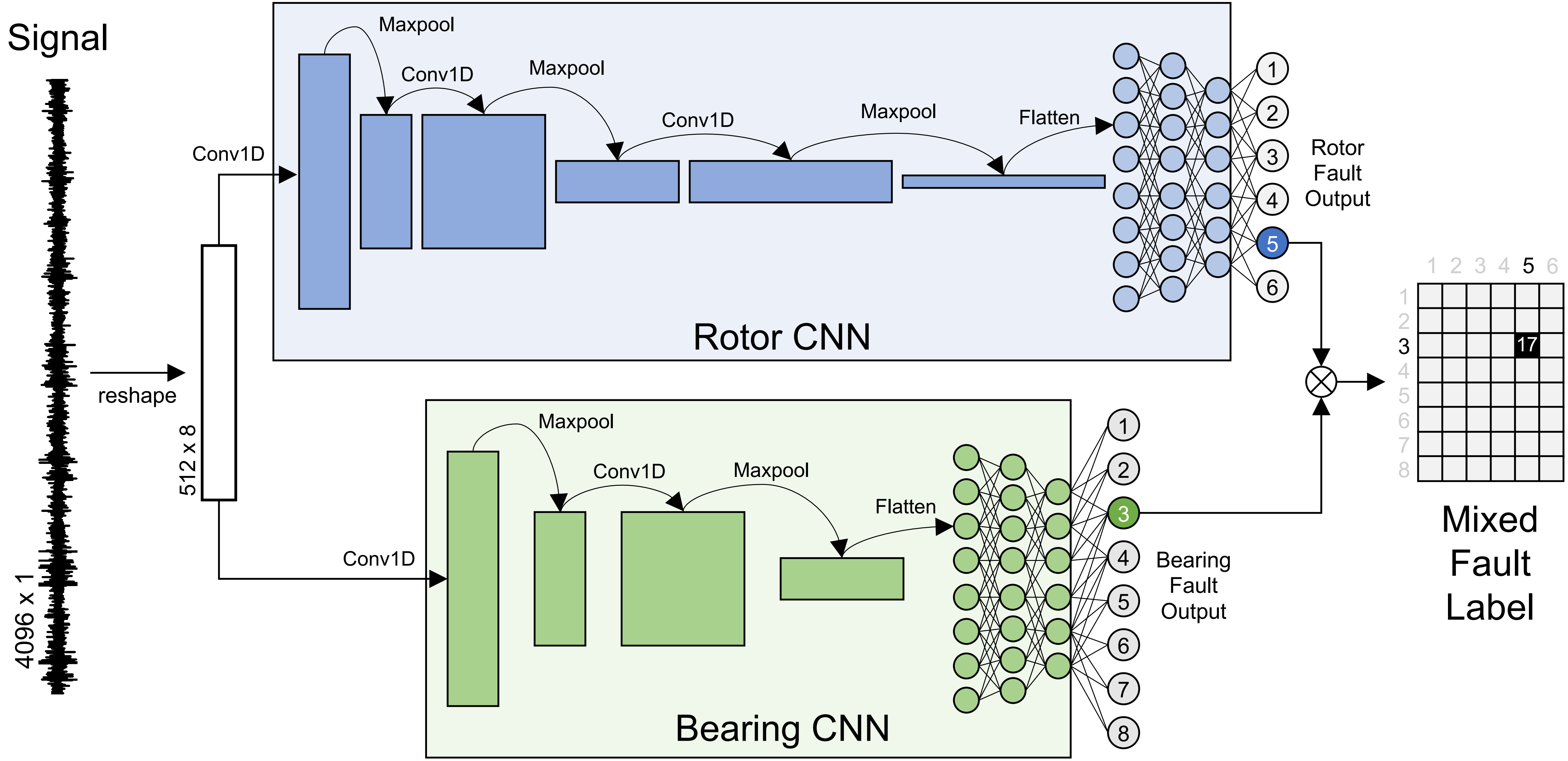
All clients have equal number of samples

Case study: Data distributions



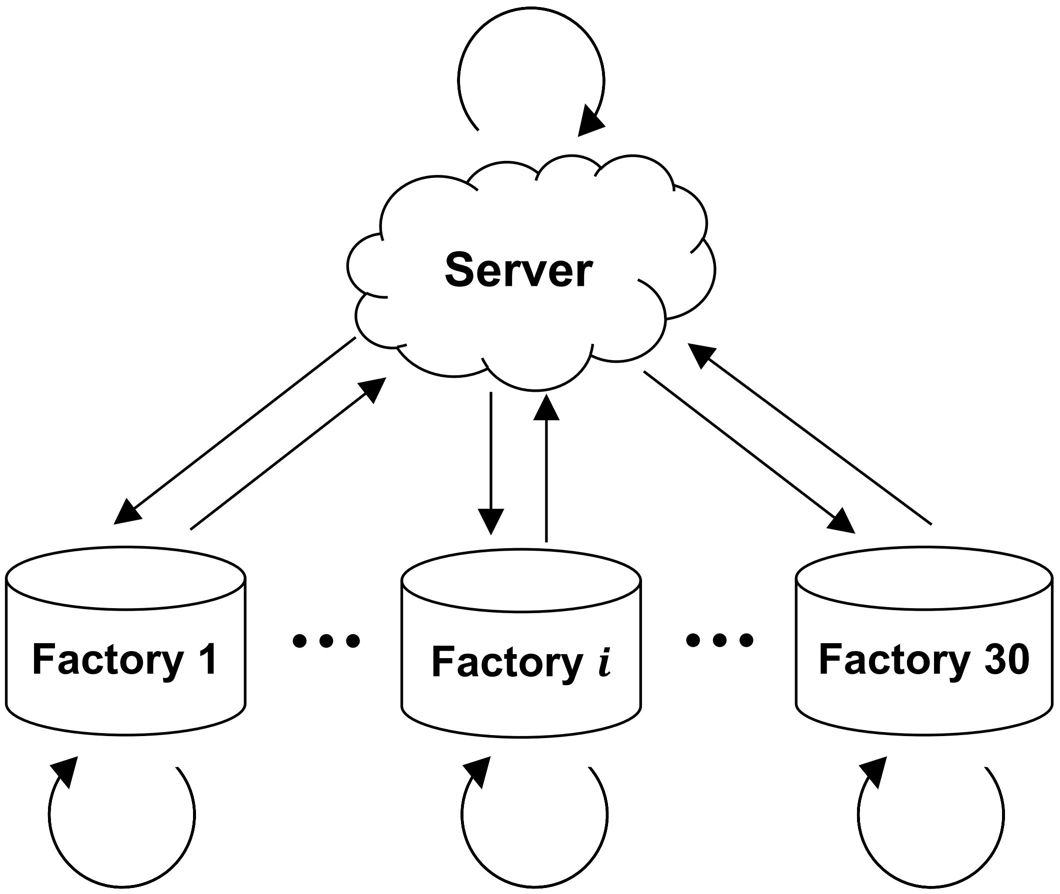
Balanced IID	Balanced non-IID	Unbalanced non-IID
All clients have all mixed fault labels	All clients do not have all mixed fault labels	All clients do not have all mixed fault labels
All clients have equal number of samples	All clients have equal number of samples	All clients do not have equal number of sample

Case study: Network architecture

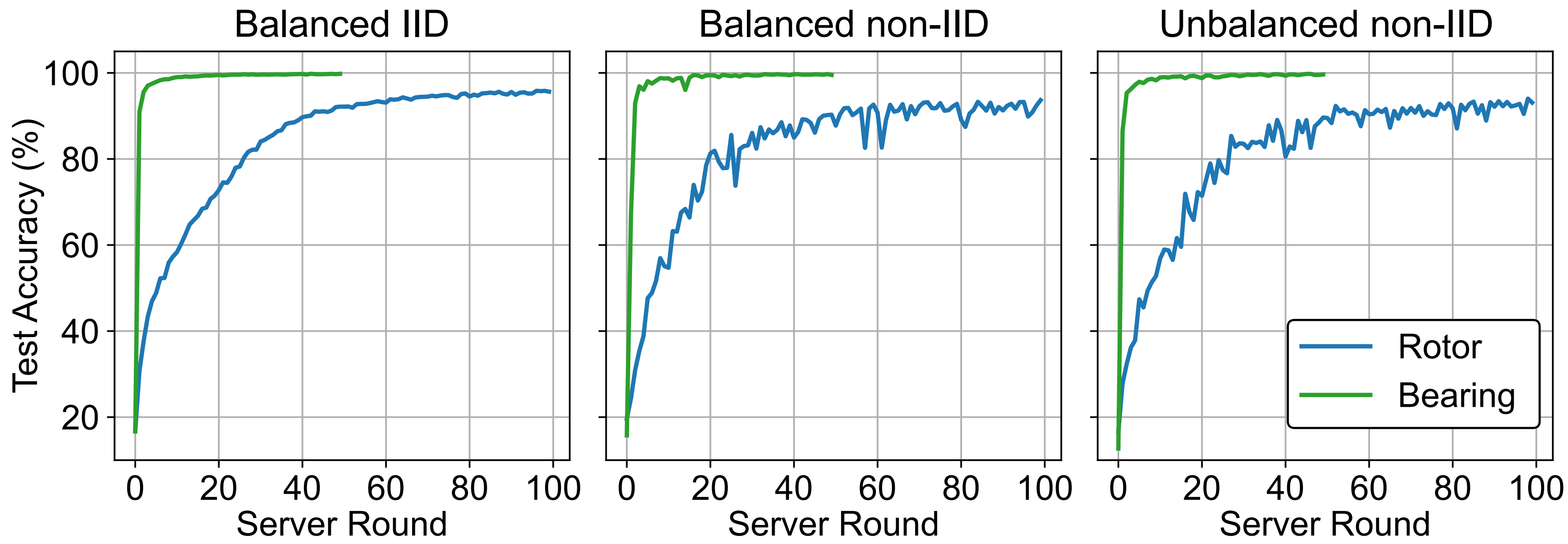


Case study: Hyperparameters

- 80-20 train-test split at each factory
- 50 server rounds for Bearing CNN
  100 server rounds for Rotor CNN
- 5 local epochs
- Client fraction 0.33 (10 out of 30 selected per round)
- Stochastic gradient descent as optimizer
- Learning rate 0.001 for all experiments



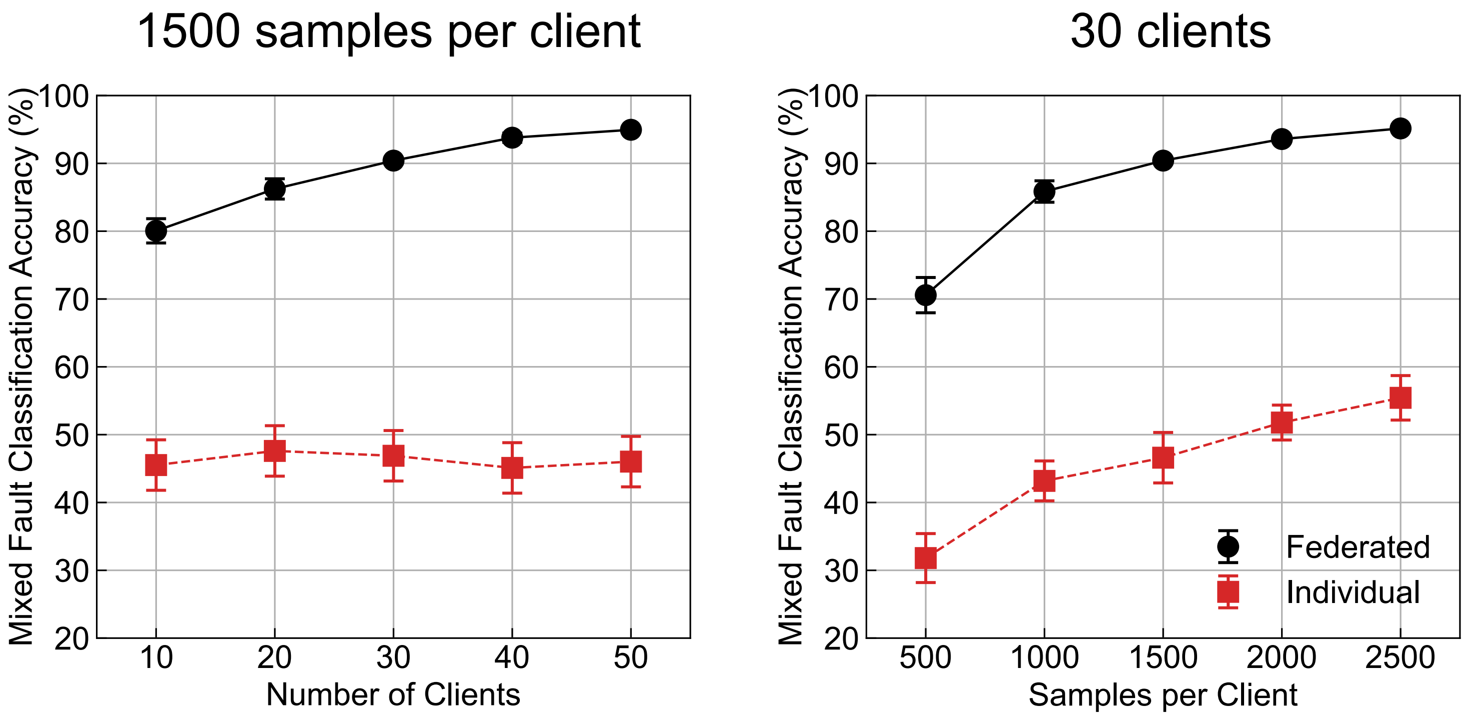
Case study: Results



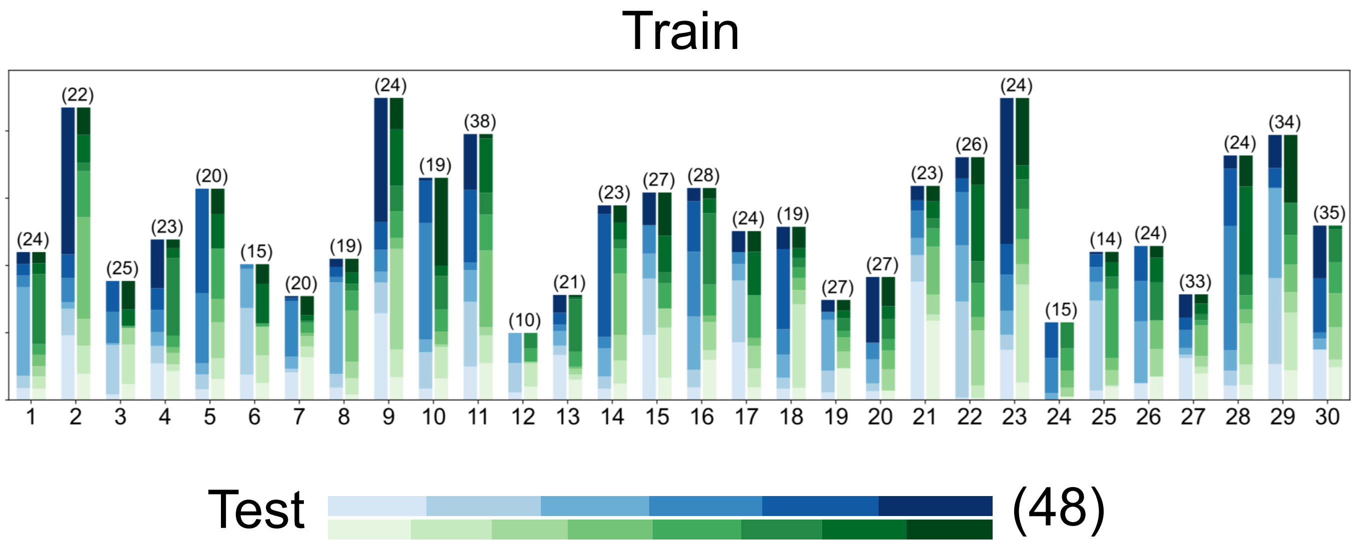
	Centralized	Federated		
		Balanced IID	Balanced non-IID	Unbalanced non-IID
Rotor	95.9	95.6	93.7	93.1
Bearing	99.8	99.7	99.6	99.5
Mixed	**95.8**	**95.5**	**93.2**	**92.8**

Performance of FL and centralized learning is comparable even for challenging data distributions

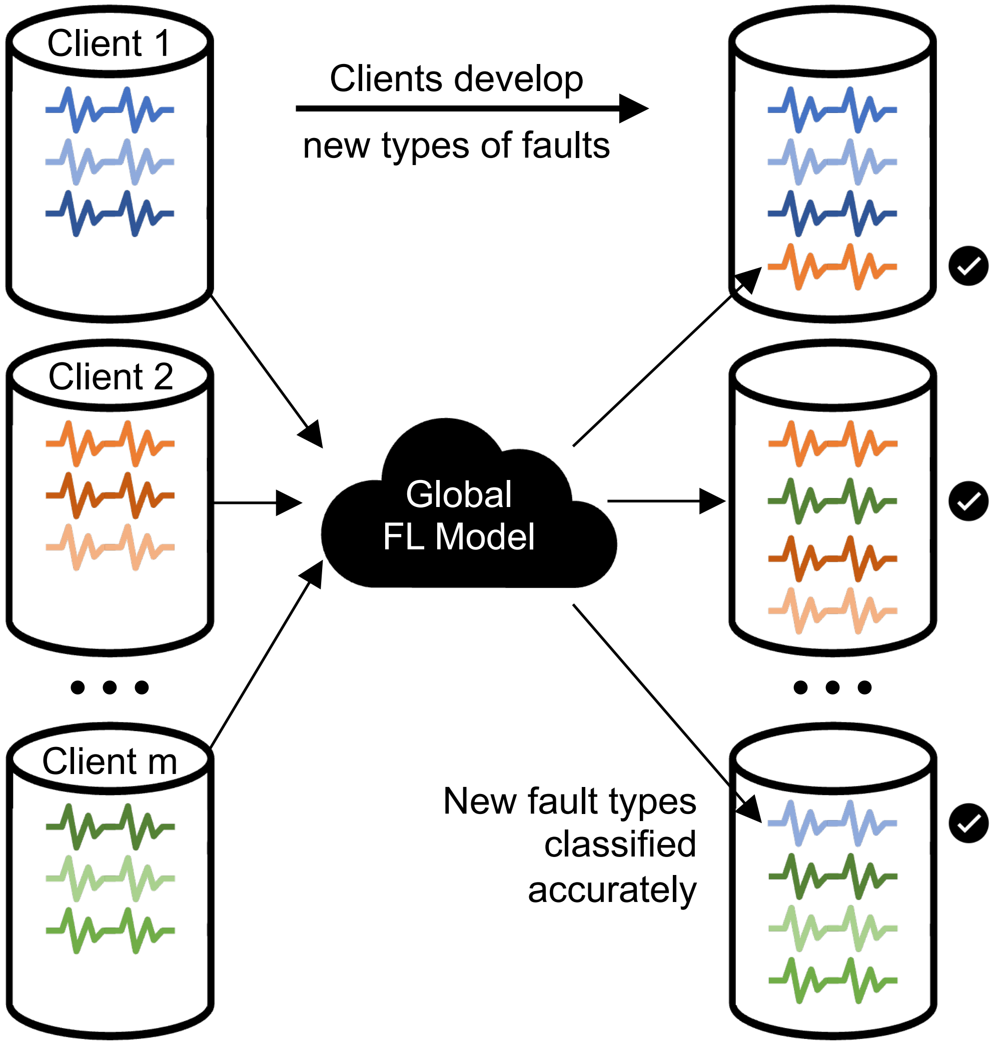
Case study: Results



- ❑ FL outperforms individual learning as factories lack sufficient data
- ❑ FL is highly data-efficient



- ❑ All clients do not have all fault types in training
- ❑ Global FL model has 92% accuracy on all 48 mixed faults
- ❑ The global FL model enables identification of previously unseen fault types



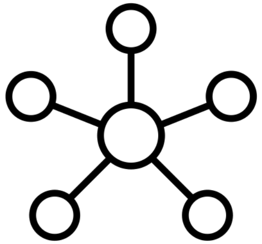
Conclusion and future work

- We propose FL-based collaborative and privacy-preserving DL for mixed fault diagnosis in rotating machinery
- FL provides a ‘win-win’ paradigm as its performance is
  - comparable to centralized learning
  - significantly better than individual learning

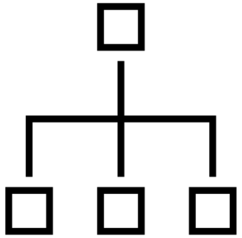
  even under unbalanced and non-IID distributions across factories

- Future work may focus on

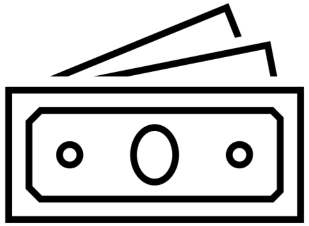
Enhanced privacy guarantees



Model personalization



System design



Sustainable incentive mechanisms

References

1. Chen, C., & Mo, C. (2004). A method for intelligent fault diagnosis of rotating machinery. *Digital Signal Processing*, *14*(3), 203-217.
2. Rafiee, J., Arvani, F., Harifi, A., & Sadeghi, M. H. (2007). Intelligent condition monitoring of a gearbox using artificial neural network. *Mechanical systems and signal processing*, *21*(4), 1746-1754.
3. Bin, G. F., Gao, J. J., Li, X. J., & Dhillon, B. S. (2012). Early fault diagnosis of rotating machinery based on wavelet packets—Empirical mode decomposition feature extraction and neural network. *Mechanical Systems and Signal Processing*, *27*, 696-711.
4. Chandra, N. H., & Sekhar, A. S. (2016). Fault detection in rotor bearing systems using time frequency techniques. *Mechanical Systems and Signal Processing*, *72*, 105-133.
5. Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., ... & Van Hoecke, S. (2016). Convolutional neural network-based fault detection for rotating machinery. *Journal of Sound and Vibration*, *377*, 331-345.
6. Xia, M., Li, T., Xu, L., Liu, L., & De Silva, C. W. (2017). Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. *IEEE/ASME transactions on mechatronics*, *23*(1), 101-110.
7. Guo, S., Yang, T., Gao, W., & Zhang, C. (2018). A novel fault diagnosis method for rotating machinery based on a convolutional neural network. *Sensors*, *18*(5), 1429.
8. Chen, S., Meng, Y., Tang, H., Tian, Y., He, N., & Shao, C. (2020). Robust deep learning-based diagnosis of mixed faults in rotating machinery. *IEEE/ASME Transactions on Mechatronics*, *25*(5), 2167-2176.
9. Yongbo, L. I., Xiaoqiang, D. U., Fangyi, W. A. N., Xianzhi, W. A. N. G., & Huangchao, Y. U. (2020). Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging. *Chinese Journal of Aeronautics*, *33*(2), 427-438.
10. Yuan, M., Wu, Y., & Lin, L. (2016, October). Fault diagnosis and remaining useful life estimation of aero engine using LSTM neural network. In *2016 IEEE international conference on aircraft utility systems (AUS)* (pp. 135-140). IEEE.
11. Yang, R., Huang, M., Lu, Q., & Zhong, M. (2018). Rotating machinery fault diagnosis using long-short-term memory recurrent neural network. *IFAC-PapersOnLine*, *51*(24), 228-232.
12. Jalayer, M., Orsenigo, C., & Vercellis, C. (2021). Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, Fast Fourier and continuous wavelet transforms. *Computers in Industry*, *125*, 103378.
13. Zhang, Y., Zhou, T., Huang, X., Cao, L., & Zhou, Q. (2021). Fault diagnosis of rotating machinery based on recurrent neural networks. *Measurement*, *171*, 108774.
14. Pei, X., Zheng, X., & Wu, J. (2021). Rotating machinery fault diagnosis through a transformer convolution network subjected to transfer learning. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-11.
15. Zhao, B., Zhang, X., Zhan, Z., & Wu, Q. (2021). Deep multi-scale adversarial network with attention: A novel domain adaptation method for intelligent fault diagnosis. *Journal of Manufacturing Systems*, *59*, 565-576.
16. Jin, Y., Hou, L., & Chen, Y. (2022). A time series transformer-based method for the rotating machinery fault diagnosis. *Neurocomputing*, *494*, 379-395.
17. Shao, H., Li, W., Cai, B., Wan, J., Xiao, Y., & Yan, S. (2023). Dual-Threshold Attention-Guided Gan and Limited Infrared Thermal Images for Rotating Machinery Fault Diagnosis Under Speed Fluctuation. *IEEE Transactions on Industrial Informatics*.

Thank You!

Manan Mehta
Department of Mechanical Science and Engineering
University of Illinois at Urbana-Champaign
mananm2@illinois.edu

Backup Slides

Manan Mehta
Department of Mechanical Science and Engineering
University of Illinois at Urbana-Champaign
mananm2@illinois.edu

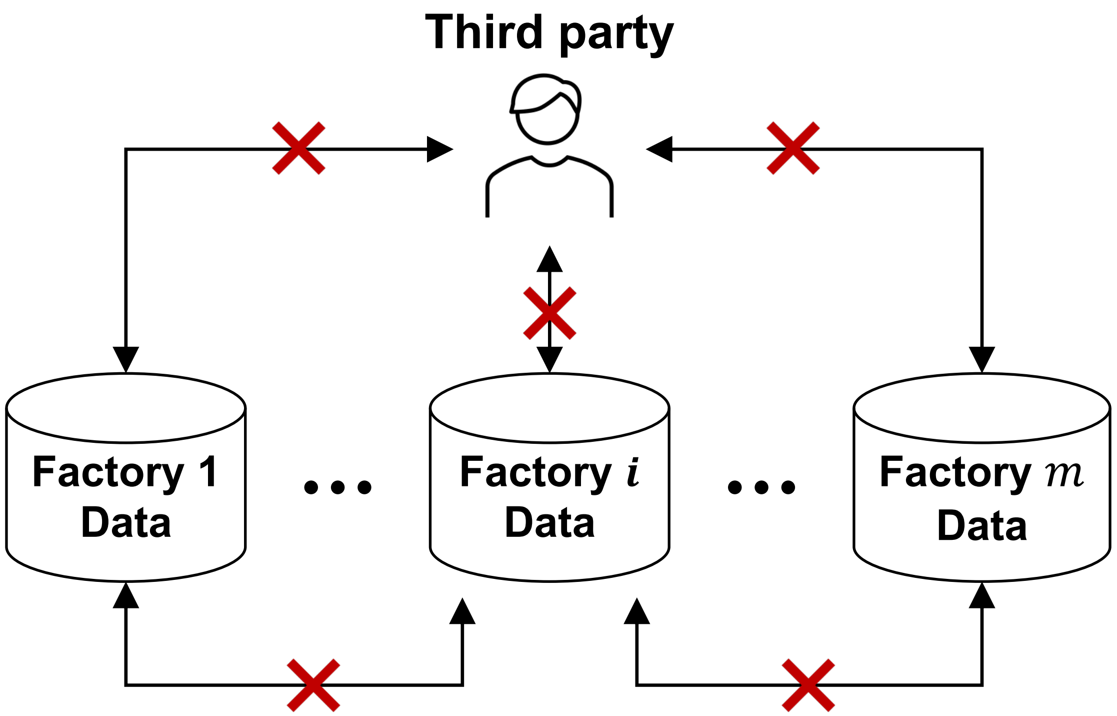
Federated learning: Problem formulation

- Fault data distributed over m factories or 'clients'
- Each factory has a local supervised learning dataset D_i of size n_i

$$D_i \coloneqq \{(x_j, y_j)\}_{j=1}^{n_i}$$

x is the vibration signal

y is the fault label



We want to construct an optimal global classifier for all m factories without directly sharing data with each other or a third party

Federated averaging

Algorithm 1 Federated Averaging (FedAvg)

Require: Distributed data $\{\mathcal{D}_k\}_1^m$ across m clients, client fraction c, number of local epochs E, local mini-batch size B, local learning rate η, number of server rounds S

Server side:

initialize the global model with parameters θ_0
for each server round $t = 1, 2, \ldots, S$ **do**
    select a random subset m_t of $c \cdot m$ clients
    send current global model θ_{t-1} to the selected clients
    for client $k \in m_t$ **in parallel do**
        $\theta_t^k \leftarrow \text{ClientUpdate}(k, \theta_{t-1})$
        receive updates θ_t^k from client k
    compute weighted average and update global model

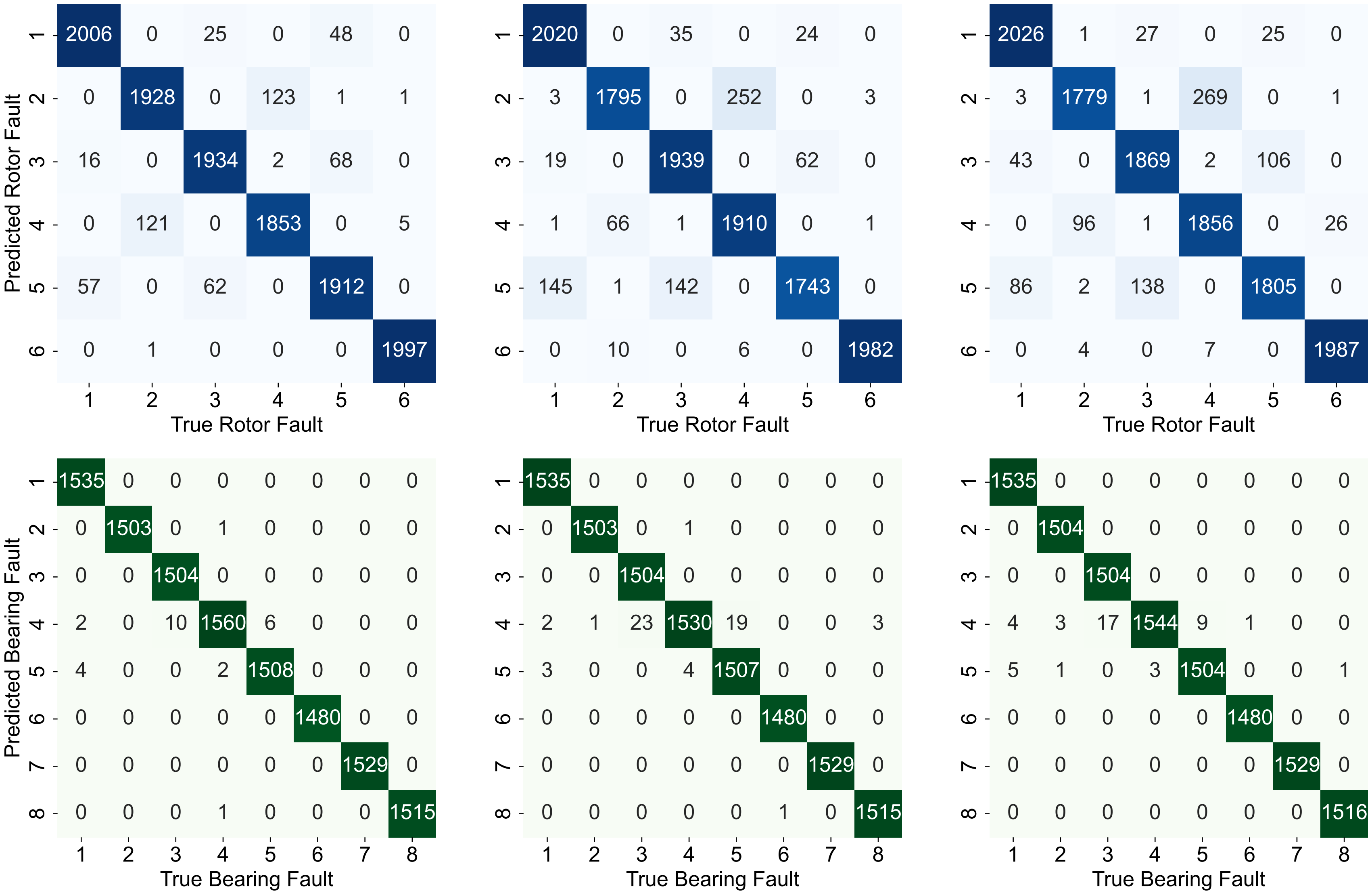
$$\theta_t \leftarrow \sum_{k=1}^{m} \frac{n_k}{n} \theta_t^k$$

Client side: Run $\text{ClientUpdate}(k, \theta)$ on client k

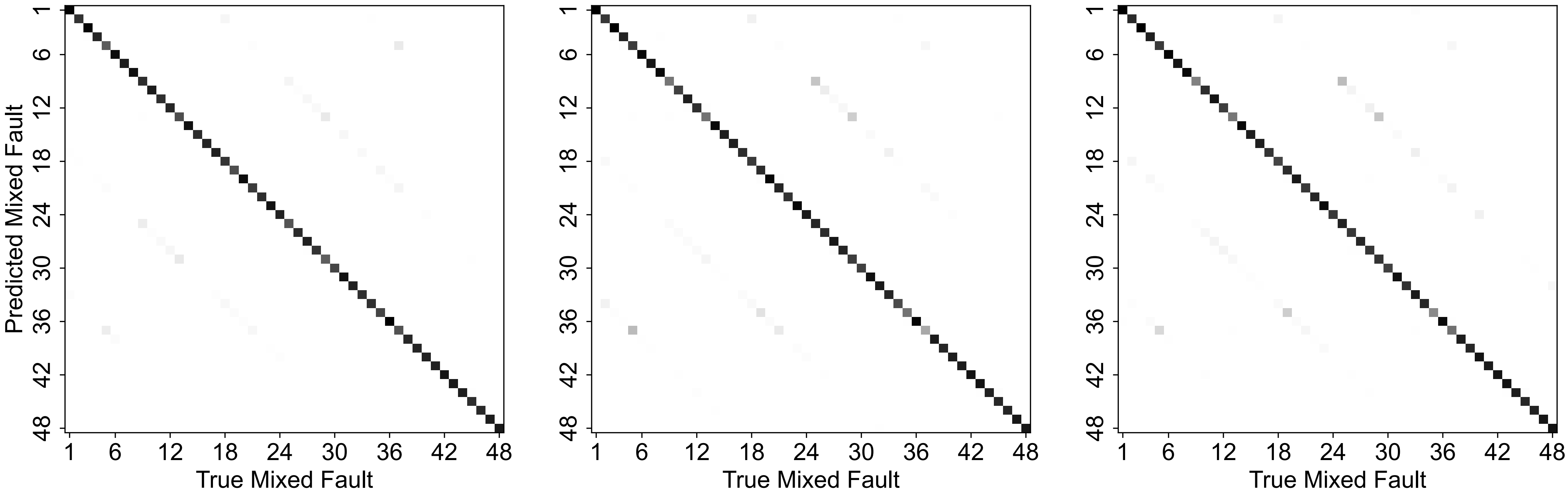
initialize the local model with θ
for each local epoch $1, 2, \ldots, E$ **do**
    for each mini-batch b of size B in $\mathcal{D}_k$ **do**
        $\theta \leftarrow \theta - \eta \nabla_\theta \ell(b)$

McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273-1282). PMLR.

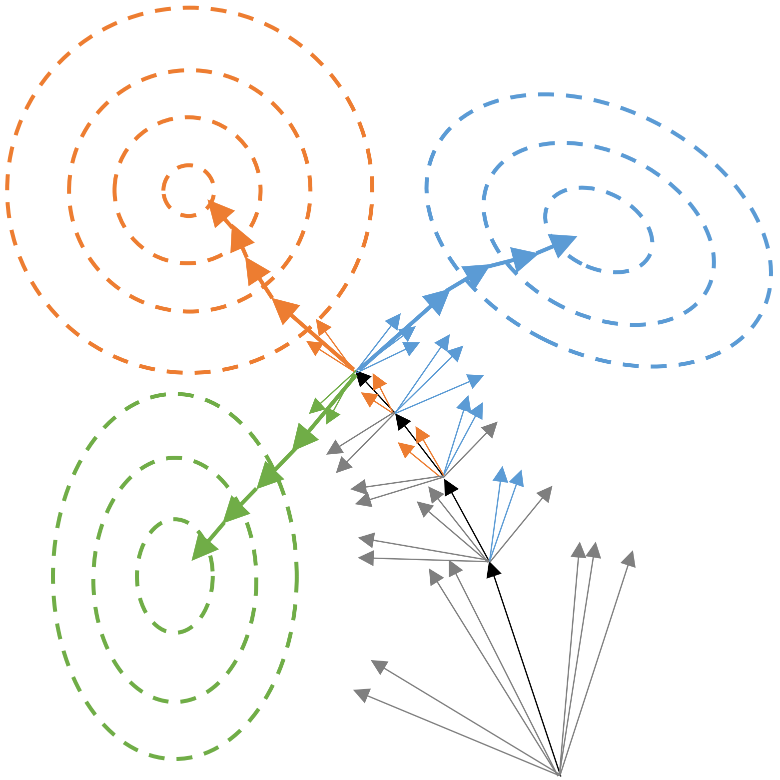
Confusion matrix



Confusion matrix

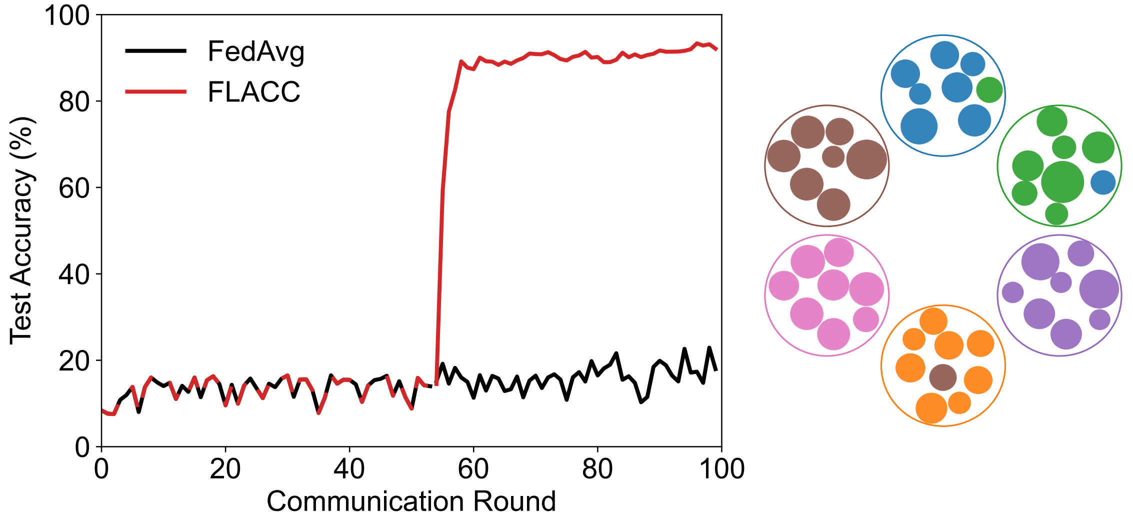


FL with extreme data heterogeneity



Federated Learning via Agglomerative Client Clustering (FLACC)

		Rotor			
		A1B1O	A1B1A	A2O	A3A
Bearing	BW1	1	2	3	4
	BW5	5	6	7	8
	OR	9	10	11	12



Mehta, M., & Shao, C. (2023). A Greedy Agglomerative Framework for Clustered Federated Learning. *IEEE Transactions on Industrial Informatics*.