Untapping Analytical Synergies in Industrial SME Ecosystems: An Empirical Evaluation of Federated Machine Learning

Anna Hensel, master's thesis

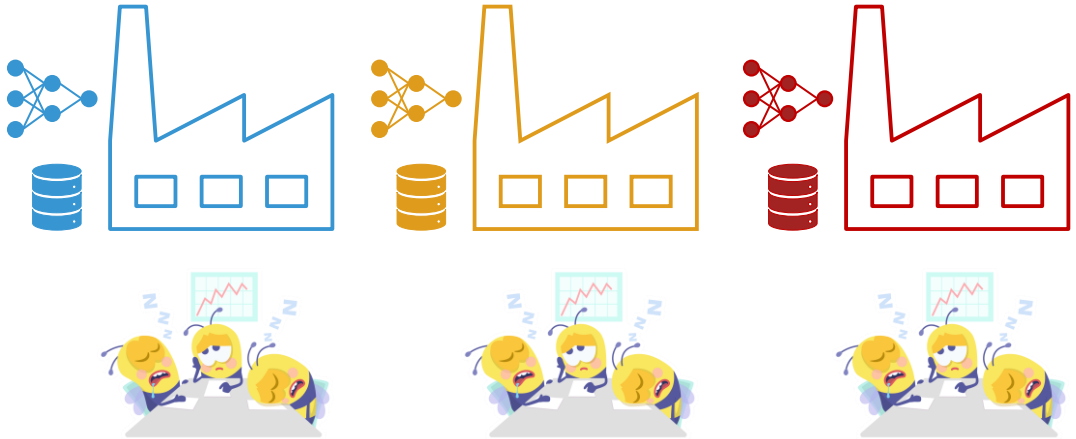
KARLSRUHE SERVICE RESEARCH INSTITUTE (KSRI)
INSTITUTE OF INFORMATION SYSTEMS AND MARKETING (IISM)

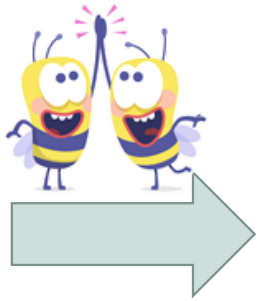
SMEs* often have related machine learning use cases, each facing several problems concerning classical ML

[7]

One model per company

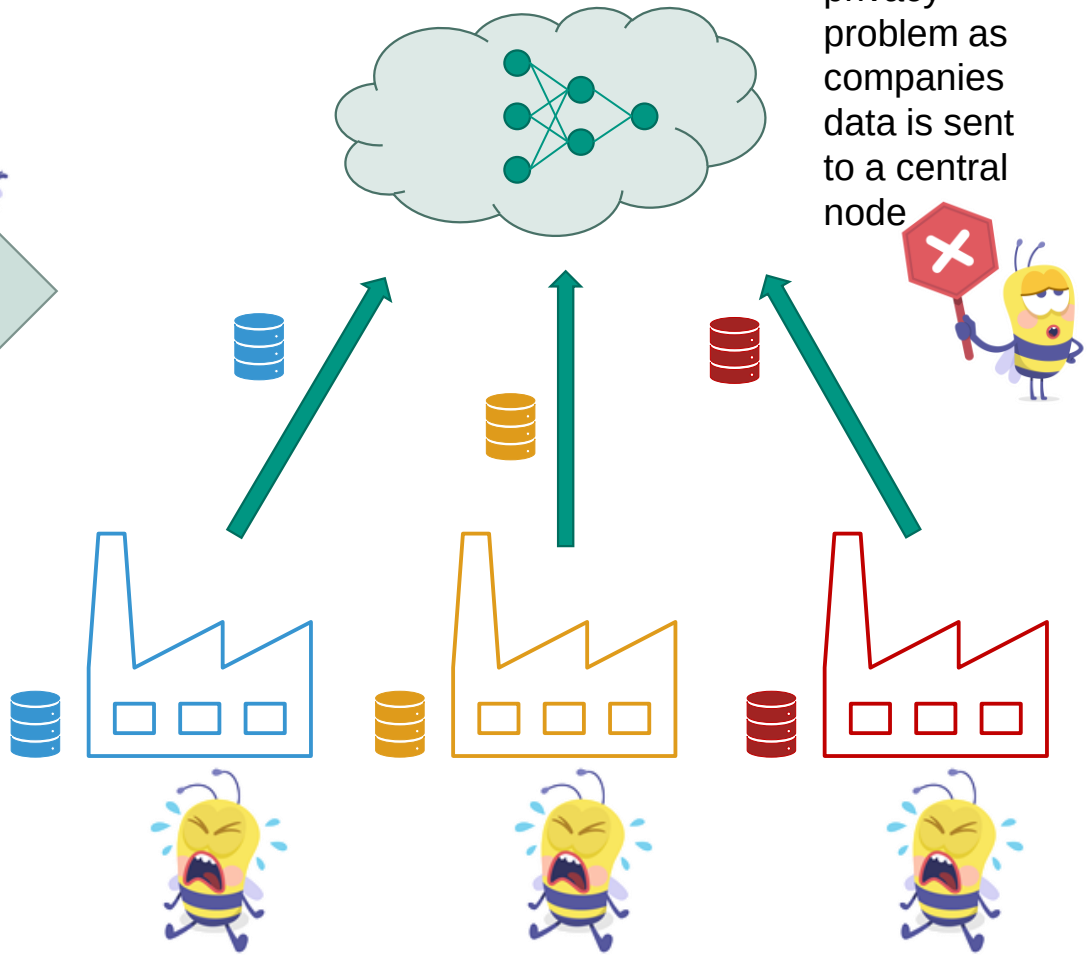
Problems: too little data, lack of experience [10, 11]
→ poor performance





All data model

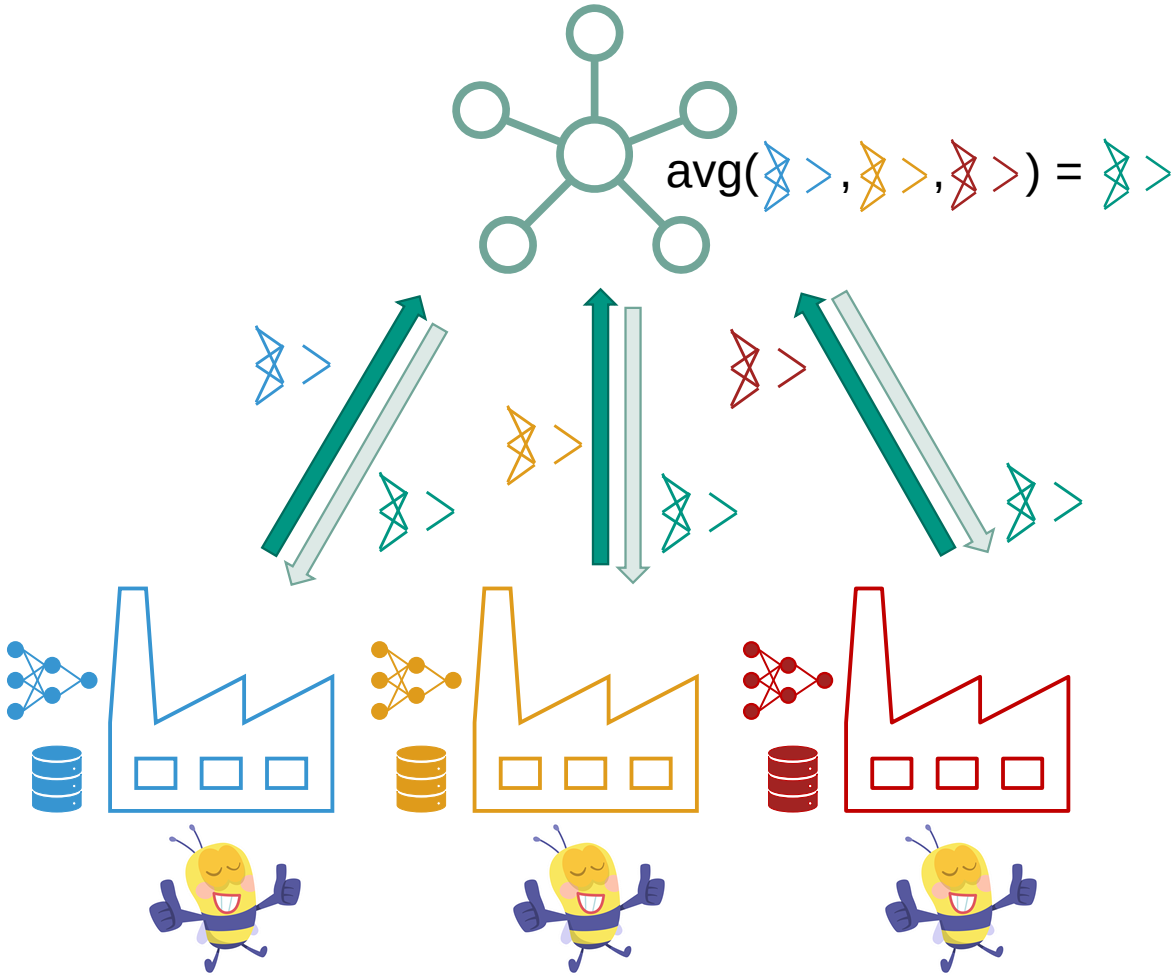
Huge privacy problem as companies data is sent to a central node



* SME: small medium-sized enterprise (dt. KMU)

Federated machine learning (FML) allows for decentralized machine learning potentially mitigating privacy issues

Federated learning model



- Local training and model sharing preserves **privacy**
- Enables joint use of data of several companies → unlocks **performance** potential
- **Complexity** due to model sharing and update communication

[2]

Agenda

Motivation

Evaluation dimensions

Privacy analysis in the SME context

Complexity analysis in the SME context

Performance analysis in the SME context

Conclusion and outlook

Privacy, complexity and performance are important dimensions for the evaluation of federated learning use cases for SMEs

Federated Learning Literature [2, 4, 5, 6, 8, 12, 13]

Privacy

- How and to what extent can private information from a specific client be reconstructed or inferred through adversarial attacks?
- Protection mechanisms

Complexity

- Computational complexity
  - Time, until specific performance is met
- Potential bottlenecks
- Speed up computation and communication

Performance

- E.g., accuracy over time / epochs
- Which FL architecture / setup performs best?

?

×

×

Privacy

- Does my data stay private?
- How can I protect my data?
- Who in the federation can be a threat?

Complexity

- Do I have the required expertise?
- How much effort is needed for organizational aspects?
  - Contracts, cost sharing, 3rd parties

Performance

- Is it worth it?

SME Perspective [10, 11]

Research Questions and Contributions

What do SMEs need to consider when evaluating FML in terms of

1. *privacy*,
2. *complexity*,
3. and *performance*?

How does FML compare to *one model per company* and *all data model*?

In summary, under which conditions is FML (not) useful for SMEs in an industrial context?

Can the knowledge and data problems be overcome using FML?

Implementation of a simulation pipeline to empirically investigate and evaluate FML in relation to *one model per company* and *all data model*.

FL aims on ensuring privacy, but there are potential privacy threats that have to be kept in mind by SMEs

Privacy

Critical aspects:
- Company sends weights to central node [2]
- Weights reflect company data

→ Adversarial attacks from “malicious client” or “malicious server” possible [8]
  - Sample reconstruction
  - Information inference
  - Model corruption
  - Runtime misclassification

Sample Reconstruction

1. Loss-Function/ReLu Exploitation
2. First Dense Layer Attack
3. DLG/iDGL (Deep Leakage from Gradients)

Not realistic / viable if:
- Clients‘ datasets sufficiently large enough
- Basic counter measures such as dropout or artificial noise

[8]

Information Inference

1. Model Inversion Attacks
2. mGAN-AI [14]
3. GAN

Not realistic / viable if:
- Input spaces sufficiently large
- Sufficient number of clients
- Non-linear models
- Heterogenous client data

[8]

Implications for SMEs

- mGAN-AI (server-side attack) as only realistic threat

→ Server has to be trustworthy

- Maybe employ 3rd party to run server / implement audits

⊕ One model per company

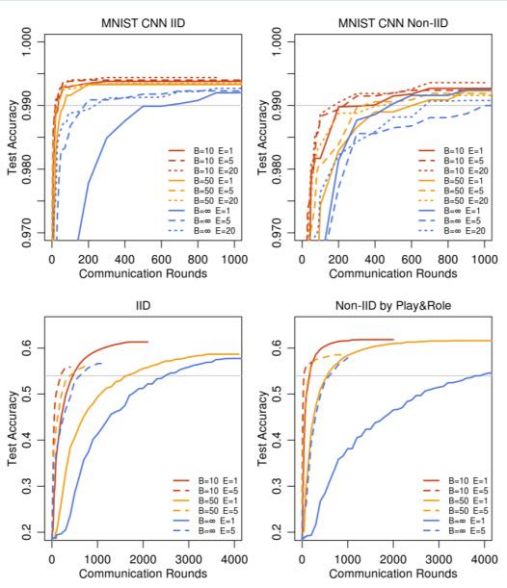
Federated learning

⊖ All data model

Complexity can be decomposed into three components which are key for evaluating federated learning in the SME context

Computational Complexity

- Number of communication rounds [2] / iterations / communicated bits until a certain target accuracy is reached [5]

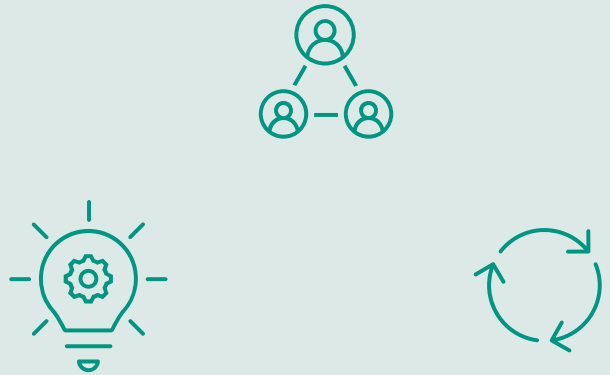


Organizational Complexity

- Complexity that arises from organizing and facilitating federated learning and the other settings, especially relating to second and third parties

Implementation Complexity

- Modelling and deployment complexity SMEs have to deal with when training ML models in the three settings



Complexity can be decomposed into three components which are key for evaluating federated learning in the SME context

Computational Complexity

- Number of communication rounds [2] / iterations / communicated bits until a certain target accuracy is reached [5]

→ In contrast to in classical federated learning use cases computational complexity is **not a dominating factor** for SMEs

Organizational Complexity

- Complexity that arises from organizing and facilitating federated learning and the other settings, especially relating to second and third parties

→ Organizational complexity, especially related to privacy, is **key driver of complexity** in facilitating federated learning in the SME context

Implementation Complexity

- Modelling and deployment complexity SMEs have to deal with when training ML models in the three settings

→ The implementation complexity for each SME in FML can be **expected to be lower** than in the one model per SME setting due to shared efforts

For analyzing the performance dimension, we implement a flexible pipeline that allows us to simulate all three settings with various parameter combinations.

Dataset

Requirements:
- real-world, industry
- publicly available
- sufficiently large
- labels / supervised

→ [7] published in June 2021 at KIT

- ~22.000 images
- Parts of spindles
- worn („pitting“) and unworn („no pitting“)

Algorithm
- CNN
- For comparability: same network architecture, optimizer, and parameters for all settings
- Early stopping routine to prevent overfitting
- Federated Averaging using TensorFlow Federated (TFF)

Performance measure to enable **comparability** between and within SMEs
- ~~Accuracy [2, 6]~~
  - ~~Threshold needed~~
  - ~~Unbalancedness~~
- AUC [9] (area under the ROC curve)
  - Abstracts from balancedness
  - Quality of score (no threshold needed)
  - Interpretation: “probability of correctly ranking a (normal, abnormal) pair” [9]

Factorial design
- *One model per client* and *all data model* setting to set the FL model into perspective
- Scenarios:

Data distribution / label distribution	Balanced	Unbalanced
Balanced	Scenario 1	Scenario 2
Unbalanced	Scenario 3	Scenario 4

  - Vary number of clients for each scenario

Data distribution / label distribution	Balanced	Unbalanced
Balanced	**Scenario 1**	Scenario 2
Unbalanced	Scenario 3	Scenario 4

Scenario 1 – balanced data distribution – balanced label distribution

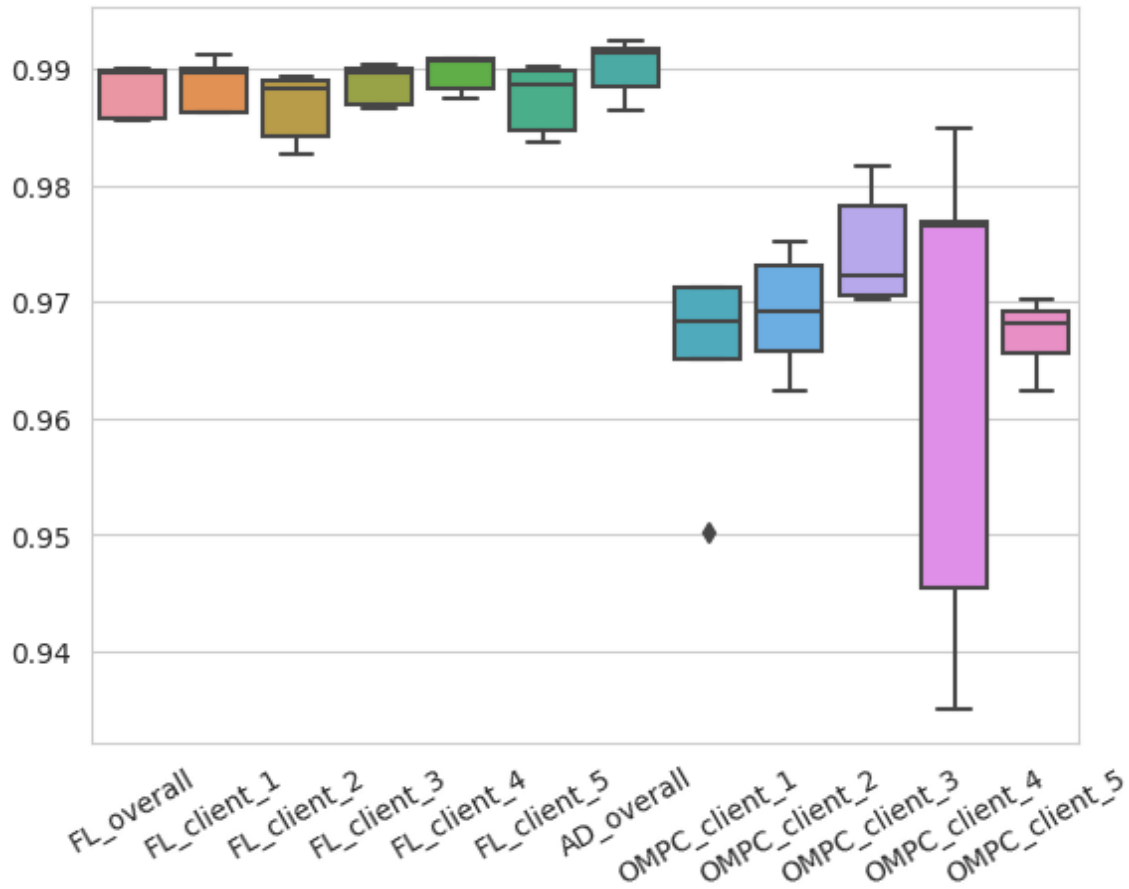
Scenario setup and resulting performance gains

client	1	2	3	4	5
weight	1	1	1	1	1
pitting share	0.5	0.5	0.5	0.5	0.5
no pitting share	0.5	0.5	0.5	0.5	0.5

client	1	2	3	4	5
performance gain	2.37%	1.97%	1.40%	2.53%	2.18%

average performance gain: 2.09%

Performance (AUC) for all settings and clients



Data distribution / label distribution	Balanced	Unbalanced
Balanced	Scenario 1	**Scenario 2**
Unbalanced	Scenario 3	Scenario 4

Scenario 2 – unbalanced data distribution – balanced label distribution

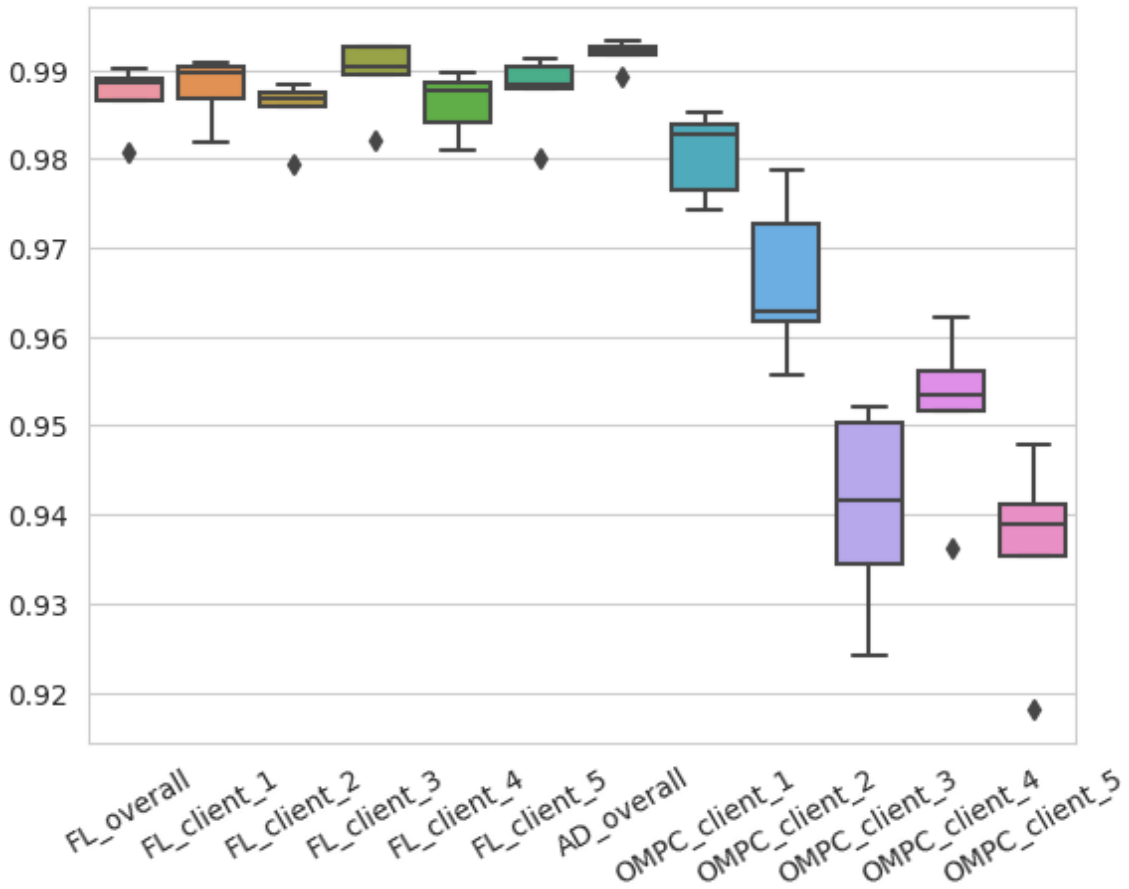
Scenario setup and resulting performance gains

client	1	2	3	4	5
weight	8	8	2	2	2
pitting share	0.5	0.5	0.5	0.5	0.5
no pitting share	0.5	0.5	0.5	0.5	0.5

client	1	2	3	4	5
performance gain	0.66%	2.13%	4.94%	3.69%	5.42%

average performance gain: 3.37%

Performance (AUC) for all settings and clients



Data distribution / label distribution	Balanced	Unbalanced
Balanced	Scenario 1	Scenario 2
Unbalanced	**Scenario 3**	Scenario 4

Scenario 3 – balanced data distribution – unbalanced label distribution

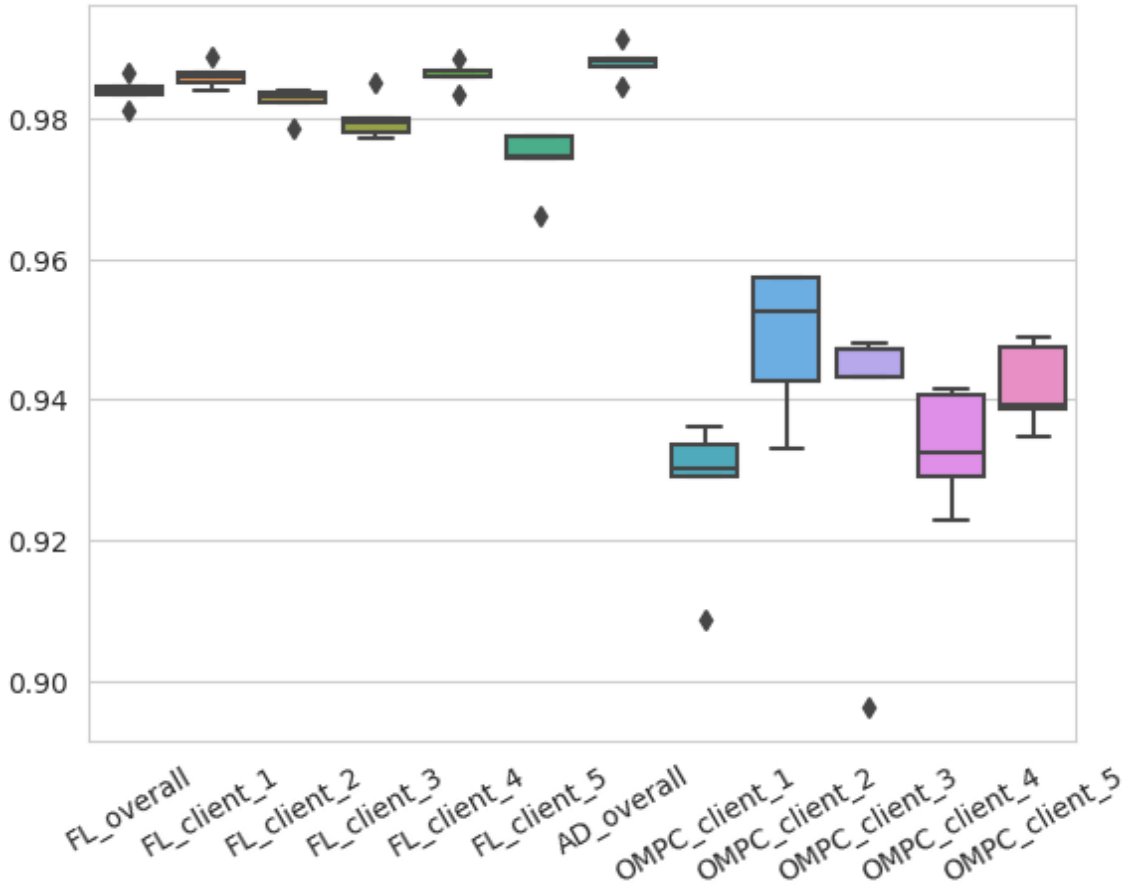
Scenario setup and resulting performance gains

client	1	2	3	4	5
weight	1	1	1	1	1
pitting share	0.8	0.2	0.8	0.2	0.8
no pitting share	0.2	0.8	0.2	0.8	0.2

client	1	2	3	4	5
performance gain	6.09%	3.72%	5.09%	5.43%	4.48%

average performance gain: 4.96%

Performance (AUC) for all settings and clients



Scenario 4 – unbalanced data distribution – unbalanced label distribution

Data distribution / label distribution	Balanced	Unbalanced
Balanced	Scenario 1	Scenario 2
Unbalanced	Scenario 3	**Scenario 4**

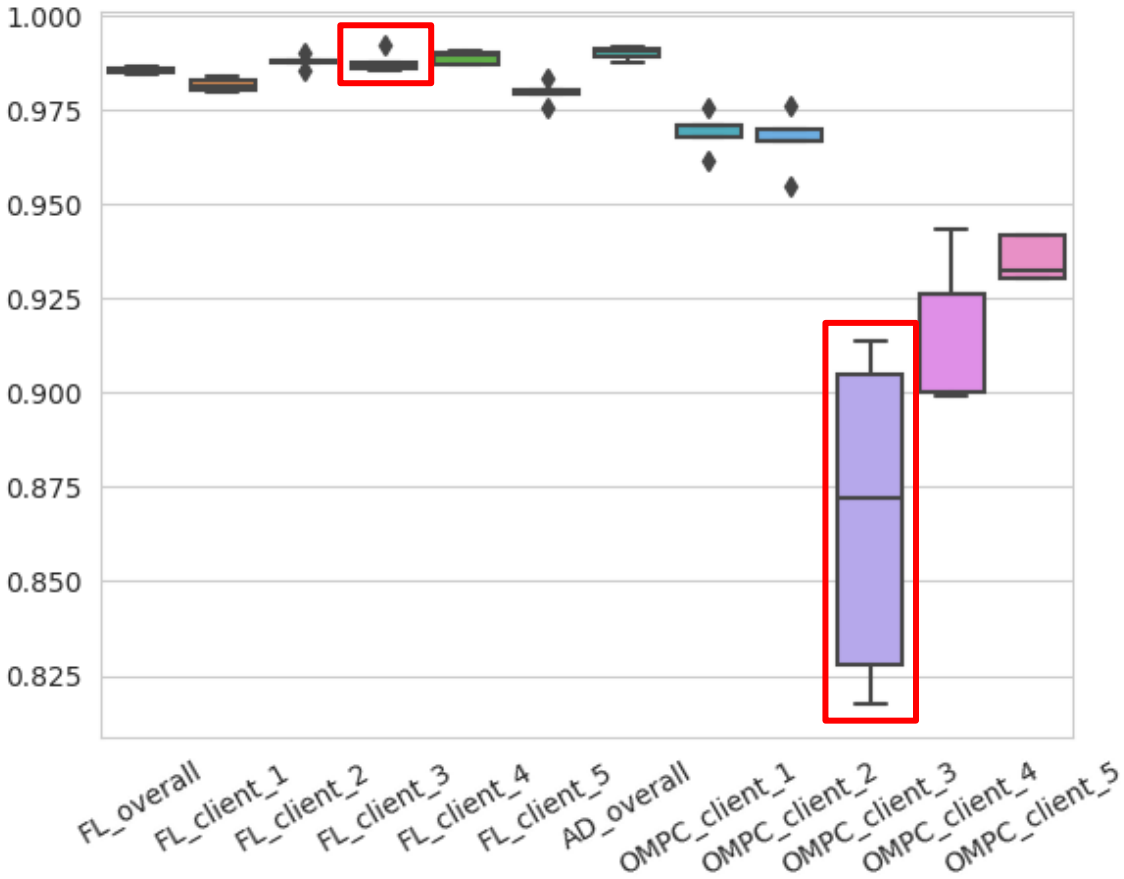
Scenario setup and resulting performance gains

client	1	2	3	4	5
weight	4	4	1	1	1
pitting share	0.8	0.2	0.8	0.2	0.8
no pitting share	0.2	0.8	0.2	0.8	0.2

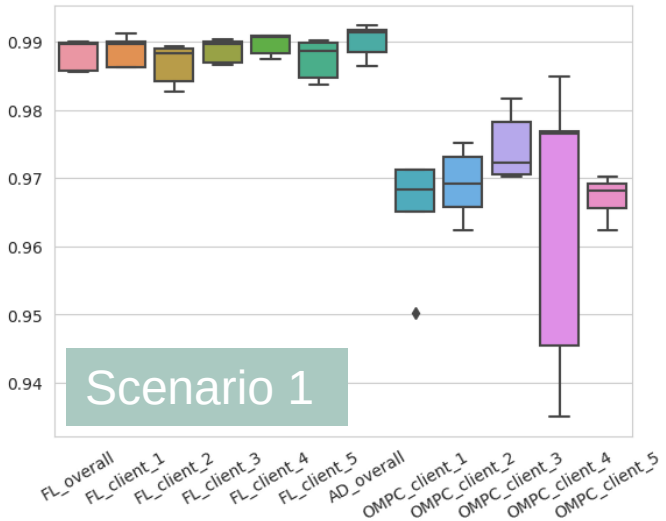
client	1	2	3	4	5
performance gain	1.67%	1.94%	13.66%	7.24%	5.37%

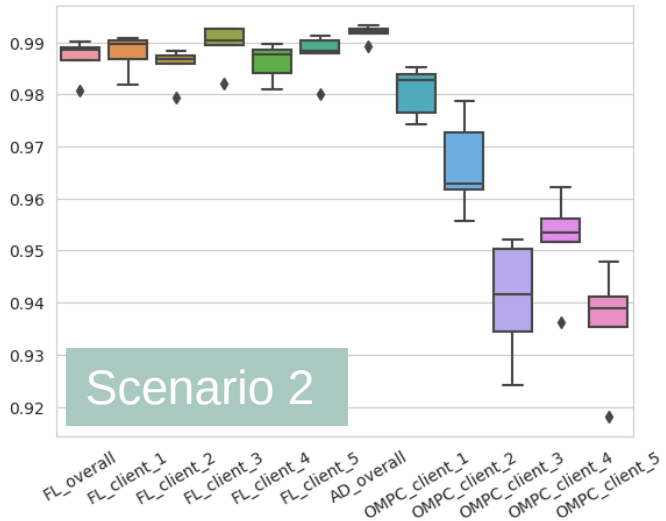
average performance gain: 5.97%

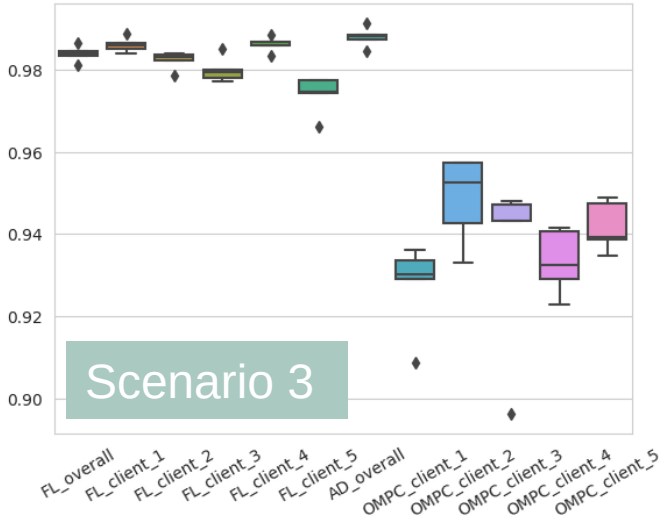
Performance (AUC) for all settings and clients

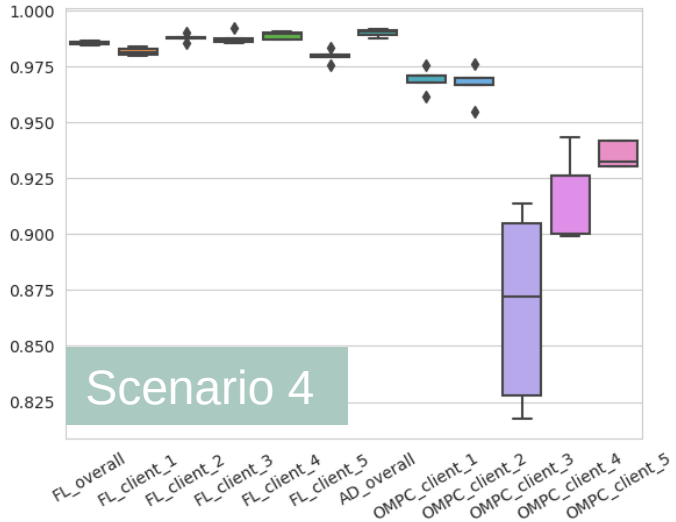


There is a clear incentive for SMEs towards taking part in federated learning from a performance perspective









- SME profits more from taking part in FL setting (measured in AUC),
  - the fewer data it has and
  - the more unbalanced the label distribution of the SME's data is
- Welfare gains of SMEs in vast majority of simulations positive → no disadvantage from taking part in FL setting even for strong SMEs
- AUC of FL models in all four scenarios at a similarly high level, even in scenarios where all SMEs face a challenging data situation

→ Clear incentive towards taking part in FL from a performance perspective

SMEs would consider taking part in FL if privacy is met, and the performance advantages outweigh the complexity disadvantages

Summary

- **Privacy**: relatively high level of privacy achievable, but can be problematic in highly sensitive settings
- **Complexity**: organizational complexity as key driver with the need for a cost sharing model that guarantees adequate incentives
- **Performance**: SME with most challenging data situation profits the most from FL which again leads to the need for an adequate incentive structure

Outlook / Limitations

- Develop organizational framework
- Develop cost sharing model
- Deep dive into knowledge and needs of SMEs

Contributions

- Evaluation of FL in an industrial SME ecosystem
- Provide guidance from a neutral point of view how the knowledge and data problems can be tackled using FL
- Simulation pipeline for Service-Meister project to demonstrate benefits of FL in various settings

Thank you for your attention!

anna.hensel@student.kit.edu

Karlsruhe Service Research Institute (KSRI)
Institute of Information Systems and Marketing (IISM)
Kaiserstr. 89 | Building 05.20
D-76133 Karlsruhe

www.ksri.kit.edu / https://dsi.iism.kit.edu/

@ksri_kit

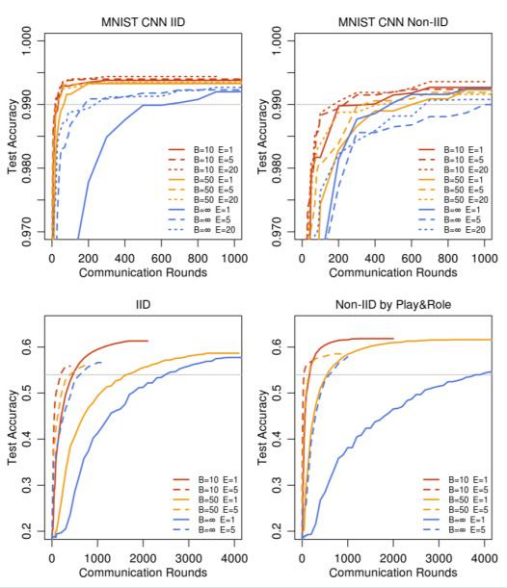
Sources

1. Bonawitz et al. 2019
2. McMahan et al. 2017
3. Murakonda et al. 2020
4. Konečný et al. 2016
5. Sattler et al. 2021
6. Yang et al. 2019
7. Schlagenhauf 2021
8. Enthoven and Al-Ars 2021
9. Hanley and McNeil 1982
10. https://hbr.org/2021/09/how-midsize-companies-can-compete-in-ai
11. https://www.eco.de/news/technischer-service-4-0-der-service-der-zukunft/
12. Li et al. 2020
13. Sun et al.
14. Wang et al. 2019

In contrast to in classical federated learning use cases computational complexity is not a dominating factor for SMEs

Computational Complexity

- Number of communication rounds [2] / iterations / communicated bits until a certain target accuracy is reached [5]



Characteristics of computational complexity per setting

Federated learning	• Cannot make ideal use of optimizers → potentially comparably slow convergence • Communication rounds can slow down training
All data model	• No restriction concerning optimizers → potentially better convergence behavior • No communication except initial data gathering and distribution of final model
One model per client	• No restriction concerning optimizers → potentially better convergence behavior • No communication

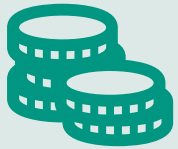
Implications for SMEs

Computational complexity not as relevant as in other use cases such as the mobile device setting → computations are not restricted to be executed on mobile devices or other specific devices with relatively low computational power and limited bandwidth and availability → manageable and not a major issue for SMEs

Organizational complexity, especially related to privacy, is key driver of complexity in facilitating federated learning in the SME context

Organizational Complexity

- Complexity that arises from organizing and facilitating federated learning and the other settings, especially relating to second and third parties



Characteristics of organizational complexity per setting

Federated learning	• Alignment of data, use case, problem formulation • Federation setup • Contract-related and legal issues • Cost-sharing model • Privacy; potential engagement of third parties for audits and a trustworthy server
All data model	• Alignment of data, use case, problem formulation • Contract-related and legal issues • Cost-sharing model
One model per client	• No organizational complexity relating to external entities

Implications for SMEs

Being pat of a FL settings comes with lots of organizational issues compared to the one model per client setting → key driver of complexity, especially related to privacy, in facilitating FL in the SME context

The implementation complexity for each SME can be expected to be lower in federated learning than in the one model per SME setting due to shared efforts

Implementation Complexity

- Modelling and deployment complexity SMEs have to deal with when training ML models in the three settings

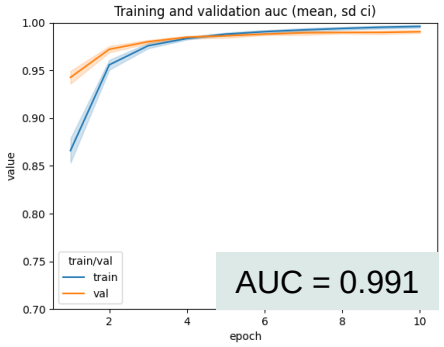
Characteristics of implemention complexity per setting

Federated learning	• Limited additional complexity for facilitating federated learning • Possibility to share implementation complexity
All data model	• Possibility to share implementation complexity
One model per client	• Each SME faces full implementation complexity

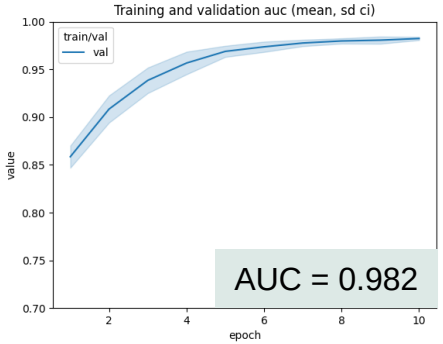
Implications for SMEs

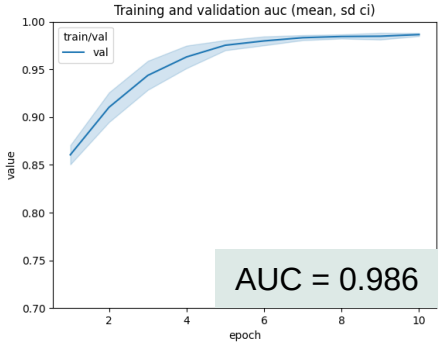
Effort, resources, and knowledge for creating a high-quality ML model could be shared.
FL frameworks such as Flower limit the additional implementation complexity due to the FL setting.
The SMEs could form a joint data science team and potentially engage a third party for support → lack of resources and knowledge could be overcome and implementation complexity could be expected to be lower in FL than in the one model per client setting.

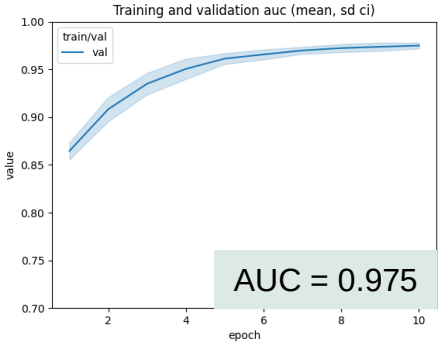
All data model



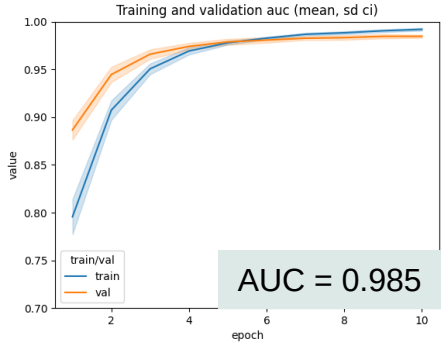
FL model per company

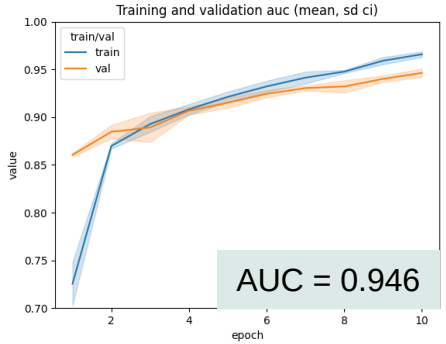


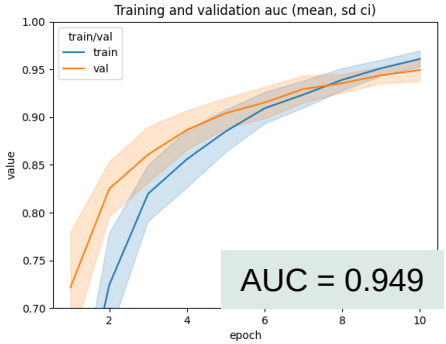




One model per company







Company 1
50% of all data, balanced

Company 2
25% of all data, unbalanced

Company 3
25% of all data, unbalanced

- The all data model can be seen as an upper bound
- One model per company can be seen as a lower bound
- Companies with insufficient data profit more from the federated model
- For companies with sufficient data there is little incentive to take part in FL setting