Privacy Preserving AI — I Federated Learning

Amogh Tarcar

May 03, 2021

Our Presenters

Bhushan Garware

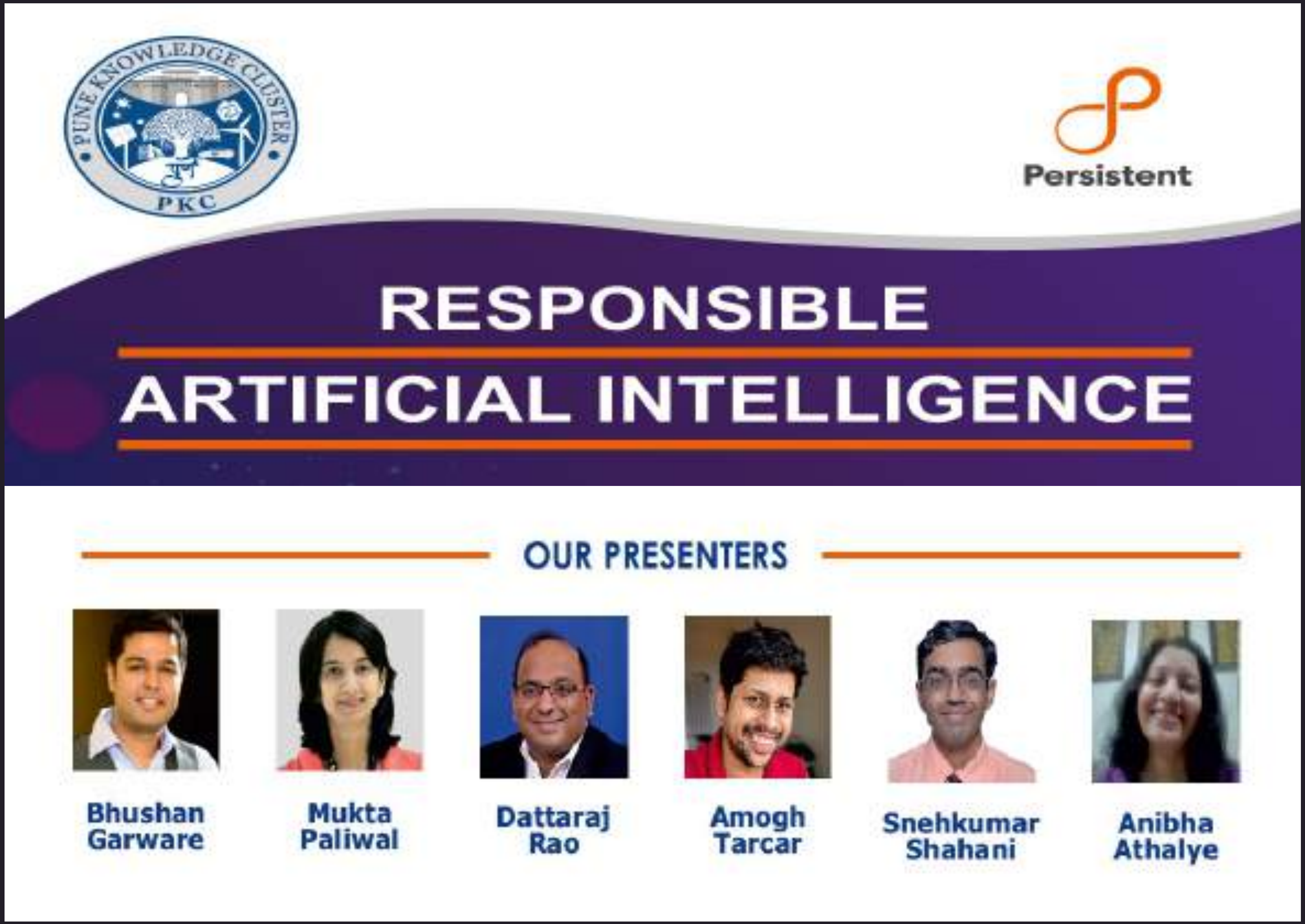
Mukta Paliwal

Dattaraj Rao

Amogh Tarcar

Snehkumar Shahani

Anibha Athalye



Course Contents :

1. Introduction to Key Pillars of Responsible AI
2. Interpretable and Explainable AI –I
3. Interpretable and Explainable AI –II
4. Fairness AI
5. **Privacy Preserving AI -I**
6. Privacy Preserving AI -II
7. Privacy Preserving AI –III
8. Secure AI
9. Reproducible AI
10. Accessible AI

Disclaimer!

Views, thoughts, and opinions expressed in this presentation belong solely to the presenter and not necessarily to the presenter's employer, organization, committee or other group or individual.

Agenda

1. Introduction to Federated Learning
2. Applications of FL
3. Demo Exercise

Introduction to Federated Learning

Responsible AI

Accountable

- Policy-driven
- Human-in-the-loop

Reproducible

- Standardized pipelines
- Data, model versioning

Transparent

- Interpretable models
- Explainable AI

Secure

- Encrypted computation
- Confidential computing

Private

- Federated learning
- Differential privacy

Startup Idea!

Build the Best Credit Card Fraud Prevention Consulting Firm in the World

Traditional ML Model Technique

Step 1

Curate labeled dataset from observed scenarios

Best representative data of the expected production scenario

Step 2

Centralize data from multiple sources at a single data center

Machine Learning requires sizable dataset to build performant models

Step 3

Centrally train a single ML model using the best ML techniques

First Step: Data!

Will the Banks Share their Data?

Hurdle 1: Banks aren't Willing to Share their Transactions Data

Start Building the Best ML Model

Startup: Smart Solution

Persistent

In Case Data cannot be Aggregated?

Challenges

- Data from a single source was inadequate for representing the real-world demographic diverse scenario
- Aggregating diverse data from multiple sources was not feasible due to legal concerns, privacy concerns or competitive market dynamics
- Centralizing data presented a potential risk of data misuse and threats concerning data security

Solution

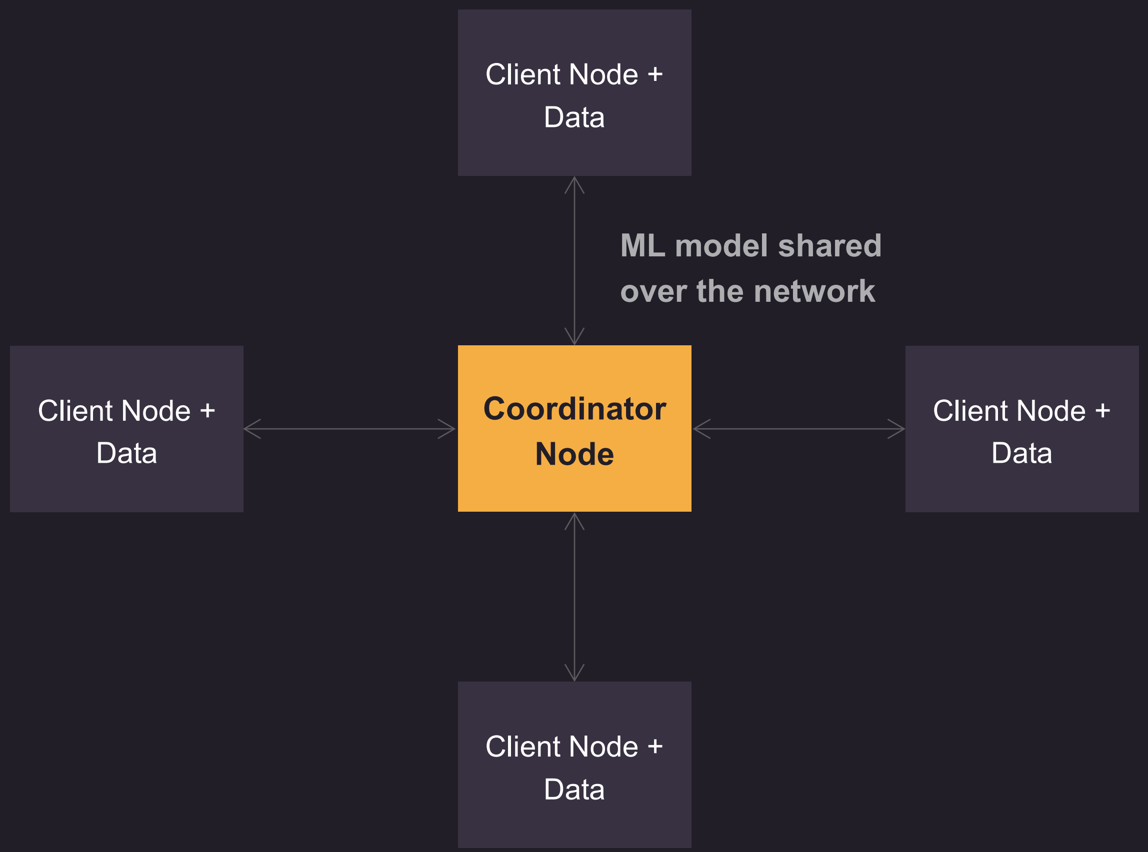
Federated Learning

A decentralized Machine Learning technique which allows multiple parties to participate in building a common global Machine Learning model, without directly sharing data.

Federated Learning

- Coordinator node orchestrates FL by starting an Active Synchronous Training Session with all the client nodes and sharing a global model
- This training session is organized in iterative steps called Training Rounds. In each training round:
  - The coordinator shares the latest version of base ML model with all the client nodes
  - Each client node runs local training on this ML model by utilizing the data present on the node and trained model updates are then shared back to the coordinator
  - Coordinator processes updates from all the nodes and fuses them together to obtain new global model
- Training rounds continue until we have a performant Global ML model with the coordinator, which is the Federated Model

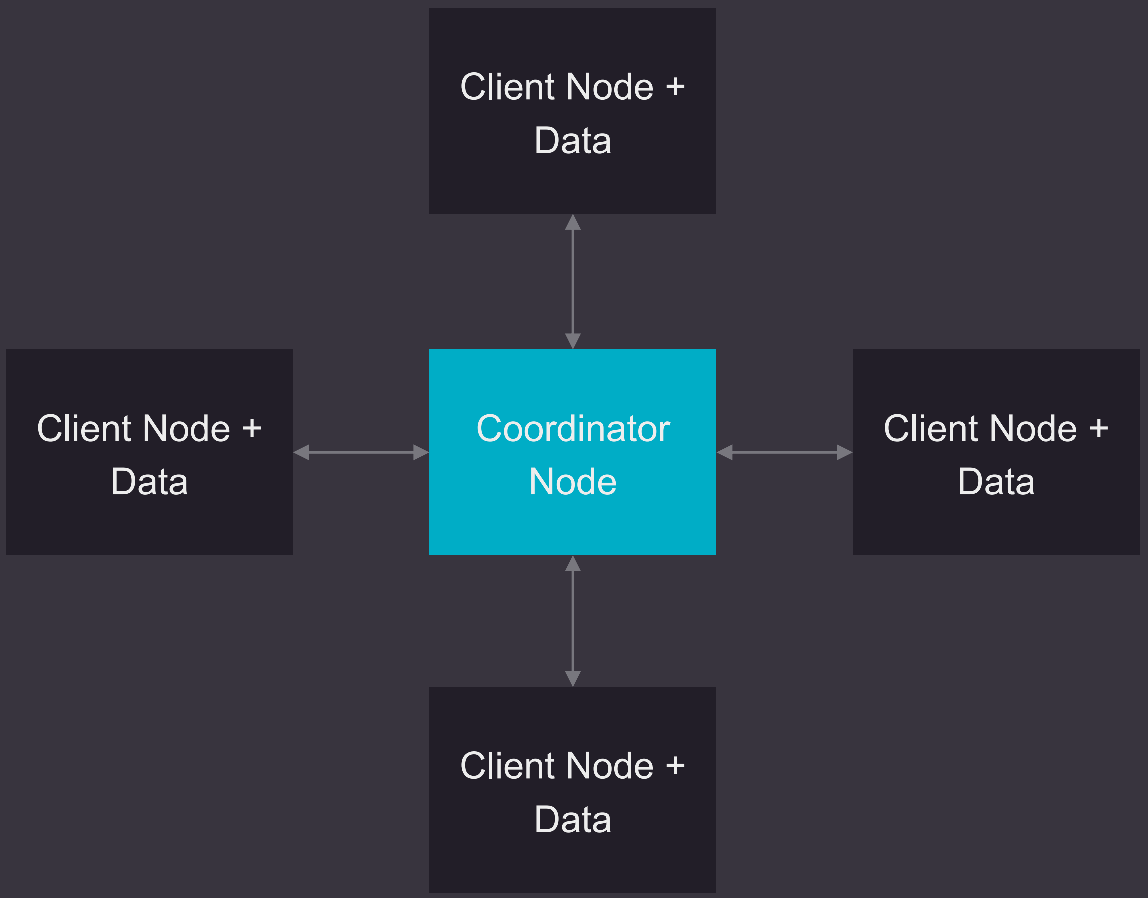
Data Never Leaves its Source Location



Federated Learning is an active area of research in Decentralized Machine Learning

Model Architectures + Fusion Algorithms

Data Never Leaves Its Source Location



Neural Networks built using Keras, PyTorch, TensorFlow

\ FedAvg, Iterative Average, Gradient Average , Fed+

Linear classifiers built using Scikit-Learn

\ Iterative Average

Decision Trees

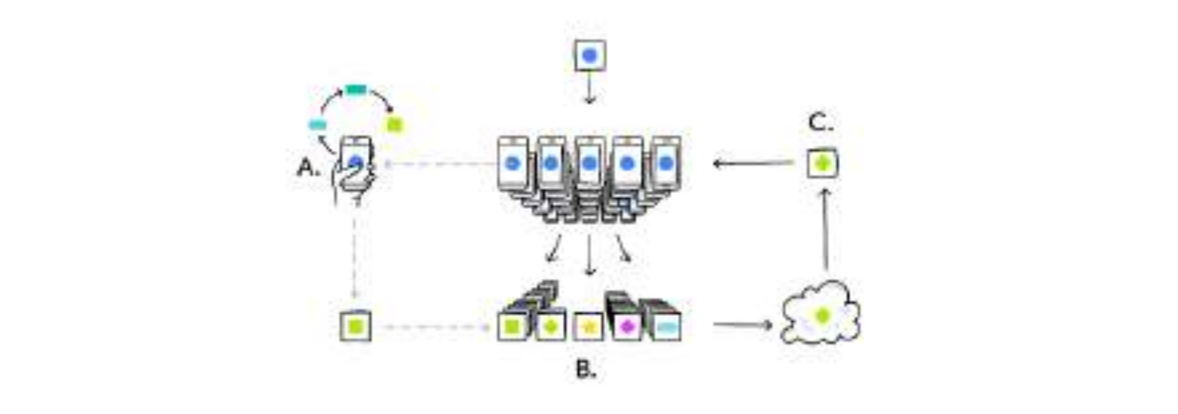
\ ID3 fusion, Federated-XGBoost

Federated Learning

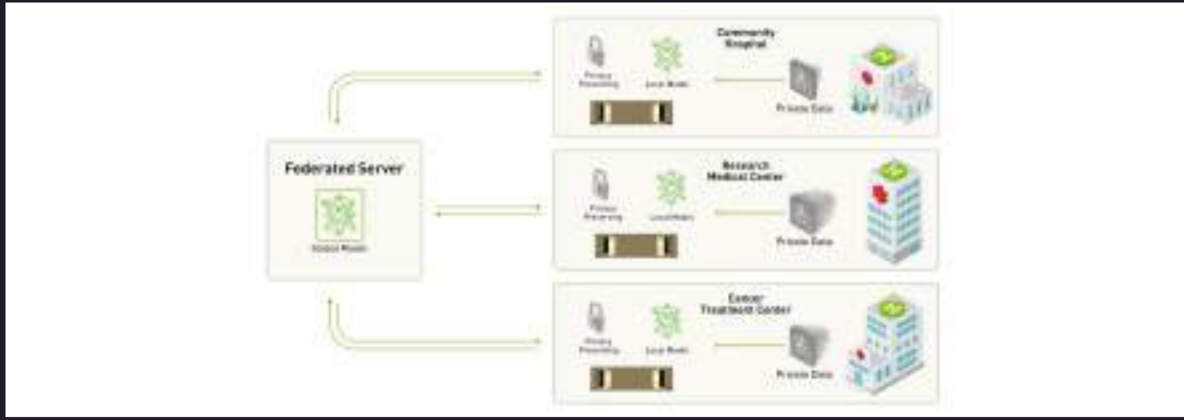
Flower Framework

NVIDIA CLARA

Federated Learning Applications

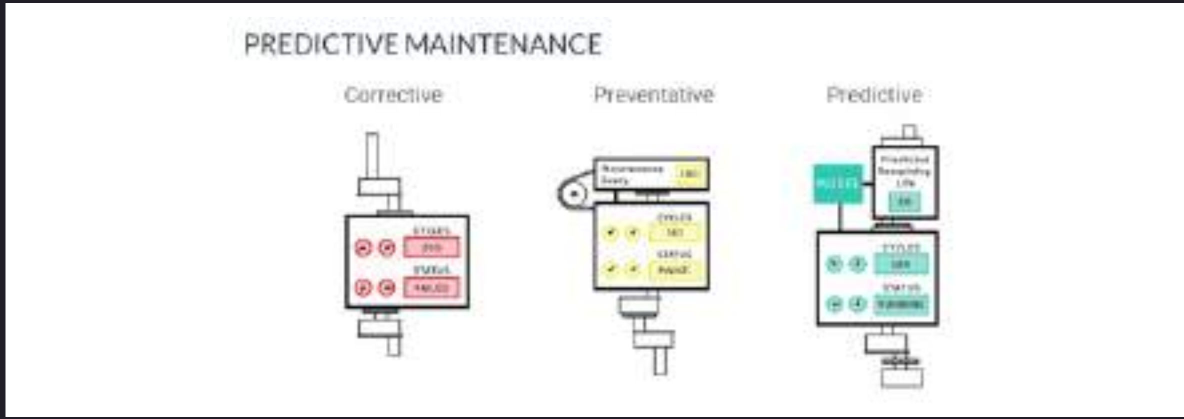


Keyboard Predictions



Healthcare Collaboration

Personalized Medicine



Preventive Maintenance

Image source: Google Federated Learning, doc.ai, nvidia clara, cloudera

Persistent

Data is the New Oil

More Data Enables Better ML Models

Smartphone Models

- Smarter typing predictions
- Pinpointed recommendations

Healthcare Models

- Diagnostic models
- Personalized medicine

Banking and Finance Models

- Fraud prevention
- Insurance/loan eligibility

Data is Siloed

- Organizational data is siloed in various business units/geographies
- Big data from IoT edge devices is expensive to aggregate

Security Concerns

- Mitigate risks of data leaks
- Robust security infrastructure in transit
- Mitigate adversarial use
- Regulatory challenges

Business Value of Data

- Diverse data is competitive advantage
- Once sold you are essentially doubling the supply

Demo

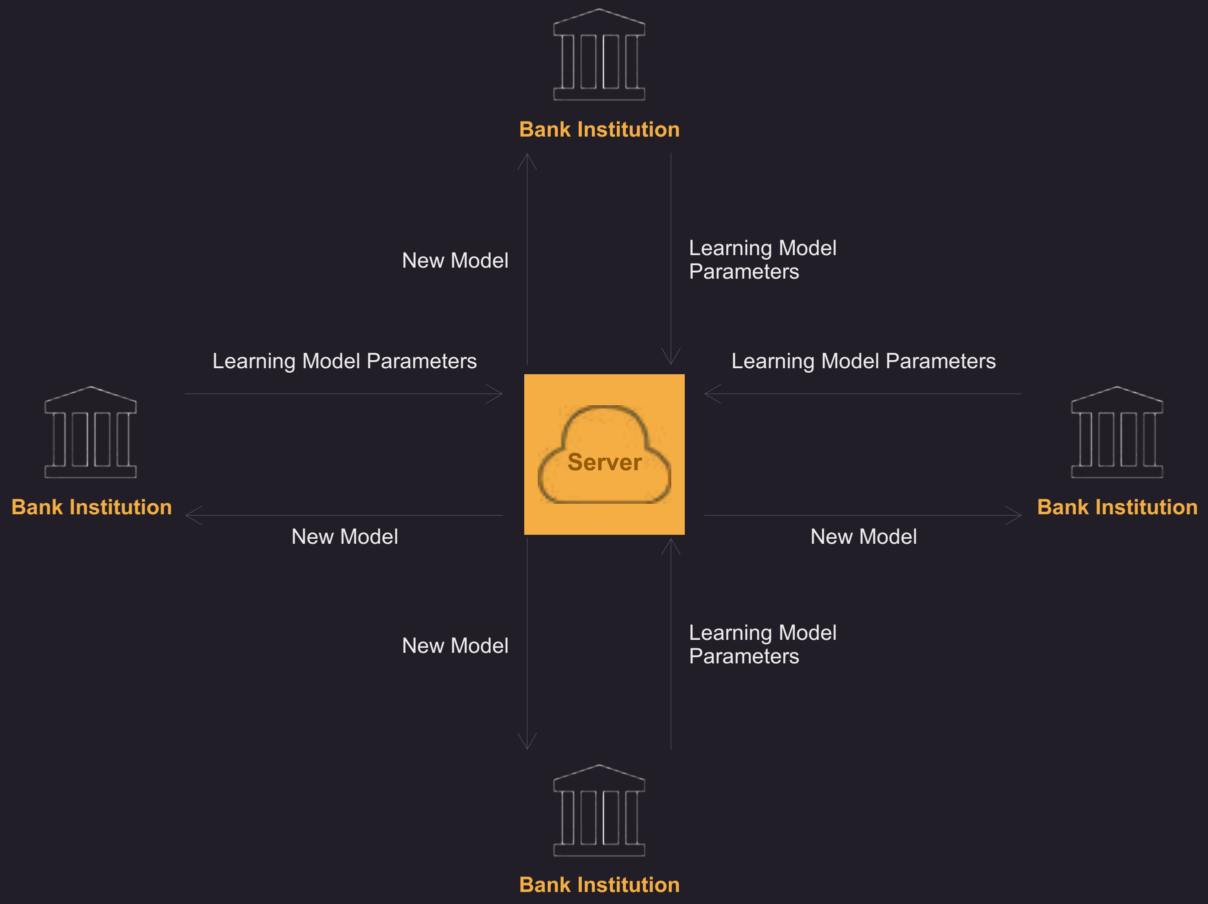
Demo of IBM FL

IBM

[Federated Learning]

Persistent

Federated Learning in Banking



Towards Federated Graph Learning for Collaborative Financial Crimes Detection

Toyotaro Suzumura[1], Yi Zhou[2], Nathalie Baracaldo[2], Guangann Ye[1], Keith Houck[1], Ryo Kawahara[3], Ali Anwar[2], Lucia Larise Stavarache[4], Yuji Watanabe[3], Pablo Loyola[3], Daniel Klyashtorny[1], Heiko Ludwig[2], and Kumar Bhaskaran[1]

[1]IBM T.J. Watson Research Center
[2]IBM Research Almaden
[3]IBM Research Tokyo
[4]IBM Global Business Services

Abstract

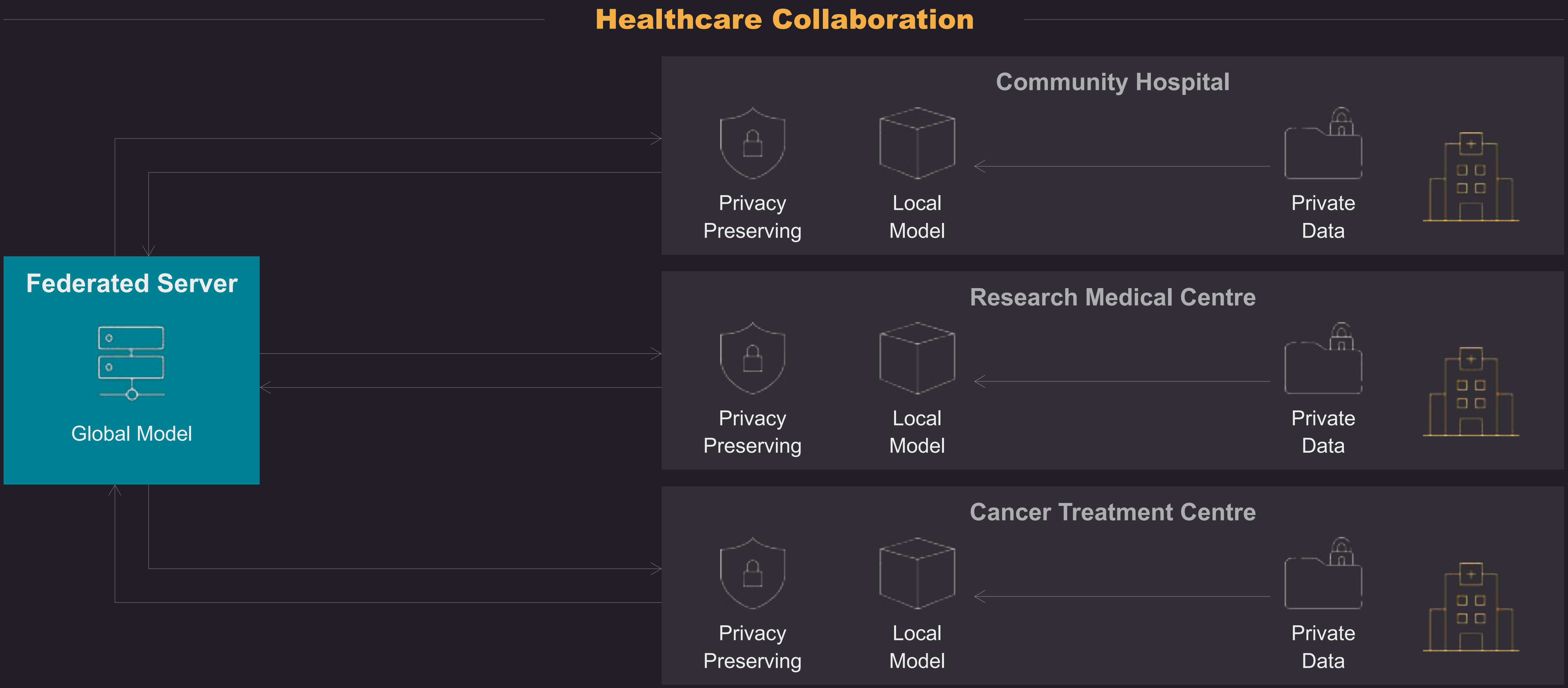
Financial crime (eg., fraud, theft, money laundering) is a large and growing problem, in some way touching almost every financial institution, as well as many individuals, and in some cases, entire societies. Financial institutions are the front line in the war against financial crime and accordingly, must devote substantial human and technology resources to this effort. Current processes to detect financial misconduct (including the technologies used) have limitations in their ability to effectively differentiate between malicious behavior and ordinary financial activity. These limitations tend to result in gross over-reporting of suspicious activity (typically manifested as "alerts") that necessitate time-intensive and costly manual review. Advances in technology used in this domain, including machine learning based approaches, can improve upon the effectiveness of financial institutions' existing processes, **however, a key challenge that most financial institutions continue to face is that they address financial crimes in isolation without any insight from other firms.** Where financial institutions address financial crimes through the lens of their own firm, perpetrators may devise sophisticated strategies that may span across institutions and geographies. Financial institutions continue to work relentlessly to advance their capabilities, forming partnerships across institutions (including governmental bodies) to share insights, patterns and capabilities. These public-private partnerships are subject to stringent regulatory and data privacy requirements, thereby making it difficult to rely on traditional technology solutions. **In this paper, we propose a methodology to share key information across institutions by using a federated graph learning platform that enables us to build more accurate machine learning models by leveraging federated learning and also graph learning approaches.** We demonstrated that our federated model outperforms local model by 20% with the UK FCA TechSprint data set. This new platform opens up a door to efficiently detecting global money laundering activity.

Source: A Federated Learning Method Used to Detect Credit Card Fraud

Paper: https://arxiv.org/pdf/1909.12946.pdf

Persistent Healthcare Experiments

Federated Learning in Healthcare



Persistent

Endoscopy Image Classification with Class Imbalance

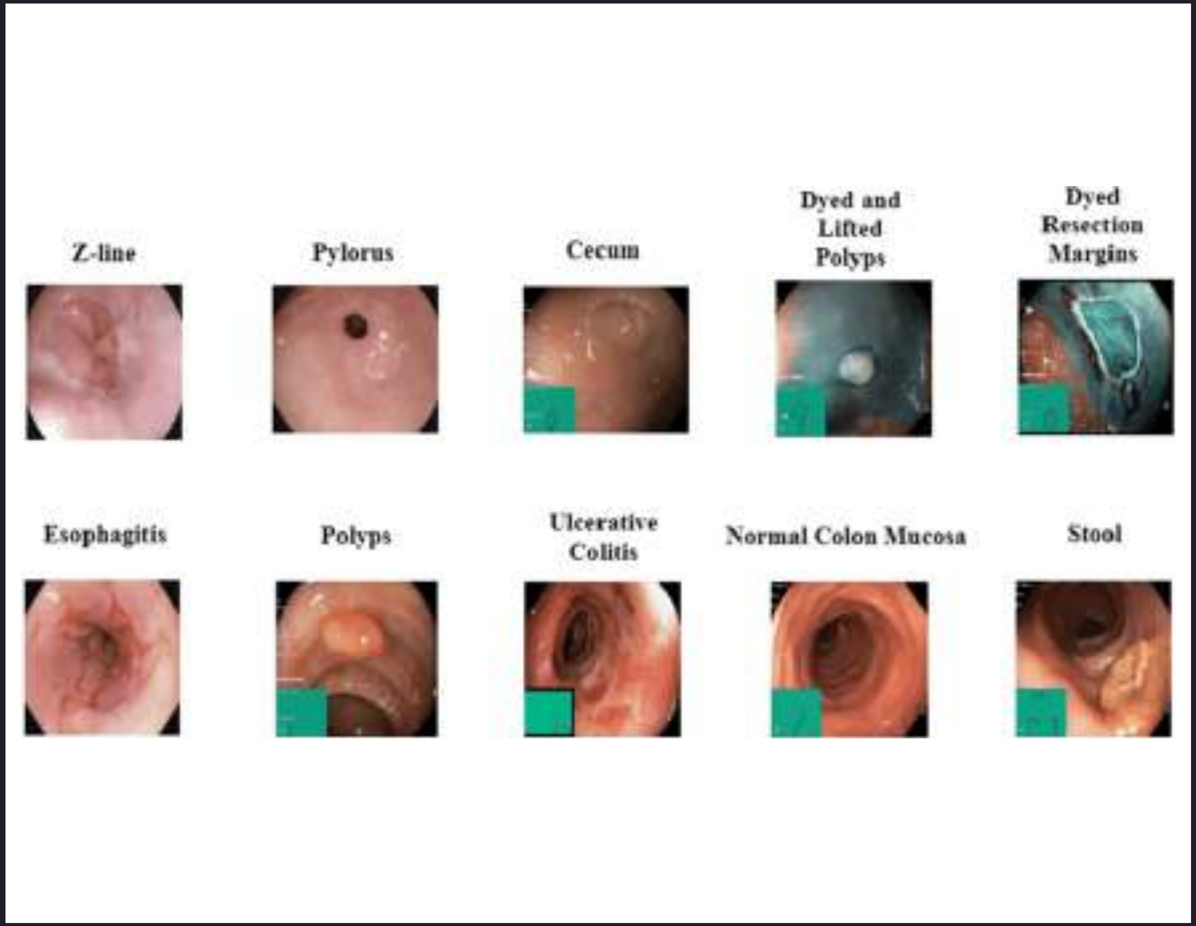
Endoscopy Images from gastrointestinal (GI) tract. Images from 8 classes labeled as anatomical and clinical findings

Split data to represent parties based on observed image resolutions

Leveraged FL frameworks for building models in a multi-party Federated Learning setting with FedAvg, Fed+ Fusion Algorithms

Federated Learning

Flower



Source: Kvasir Paper

Persistent

Endoscopy Image Classification with Class Imbalance

Imbalanced Class Distributions across Parties

Party_576	
Class 0	822
Class 1	784
Class 2	762
Class 3	57
Class 4	27
Class 5	772
Class 6	808
Class 7	78
Total Samples	**4110**

Party_1024	
Class 2	5
Class 3	567
Class 4	583
Class 5	7
Class 7	535
Total Samples	**1697**

Party_1072	
Class 1	2
Class 2	14
Class 3	164
Class 4	188
Class 5	15
Class 6	1
Class 7	203
Total Samples	**587**

Class 0	'normal-cecum'
Class 1	'dyed-resection-margins'
Class 2	'ulcerative-colitis
Class 3	'esophagitis'
Class 4	'normal-pylorus'
Class 5	'polyps'
Class 6	'dyed-lifted-polyps'
Class 7	'normal-z-line'

	Dominant Resolution Was	Total Training Samples
Party 1	576p	**4110**
Party 2	1024p	**1697**
Party 3	1072p	**587**

Total Training Samples: ~6400

Testing Samples: 1600

Dominant Resolutions across Classes Example

Polyps	576	969
Polyps	1072	20
Polyps	1024	10
Polyps	1064	1
Dyed-lifted-polyps	576	996
Dyed-lifted-polyps	1080	1
Dyed-lifted-polyps	1064	2
Dyed-lifted-polyps	1072	1

Ulcerative-colitis	576	974
Ulcerative-colitis	1072	19
Ulcerative-colitis	1024	5
Ulcerative-colitis	1080	2
Normal-pylorus	1024	730
Normal-pylorus	1072	237
Normal-pylorus	576	33
Normal-cecum	576	1000

Normal-z-line	1024	649
Normal-z-line	576	102
Normal-z-line	1072	249
Esophagitis	1072	206
Esophagitis	1024	728
Esophagitis	576	66
Dyed-resection-margins	576	998
Dyed-resection-margins	1072	2

Class	Resolution	No. of Images
Polyps	**576*720**	**969 <- Dominant**
	1072*1920	20
	1024*1280	10
Esophagitis	576*720	66
	1072*1920	206
	1024*1280	**728 <- Dominant**

Class	Resolution	No. of Images
Normal-z-line	576*720	102
	1072*1920	249
	1024*1280	**649 <- Dominant**

Endoscopy Image Classification with Class Imbalance Results

Classification Report

	Precision	Recall	F1-score	Support
0	0.82	0.98	0.89	178
1	0.86	0.91	0.88	214
2	0.95	0.84	0.89	217
3	0.77	0.82	0.79	212
4	0.93	0.94	0.93	202
5	0.89	0.84	0.87	205
6	0.88	0.78	0.83	188
7	0.75	0.71	0.73	184
Accuracy			0.85	1600
Macro Avg	0.86	0.85	0.85	1600
Weighted Avg	0.86	0.85	0.85	1600

	0	1	2	3	4	5	6	7
0	175	0	0	0	1	2	0	0
1	0	194	0	0	1	0	19	0
2	16	1	183	2	6	8	0	1
3	0	0	0	174	0	0	0	38
4	0	0	2	2	190	4	0	4
5	19	0	4	1	7	173	1	0
6	4	31	1	0	0	5	147	0
7	0	0	2	48	0	3	0	131

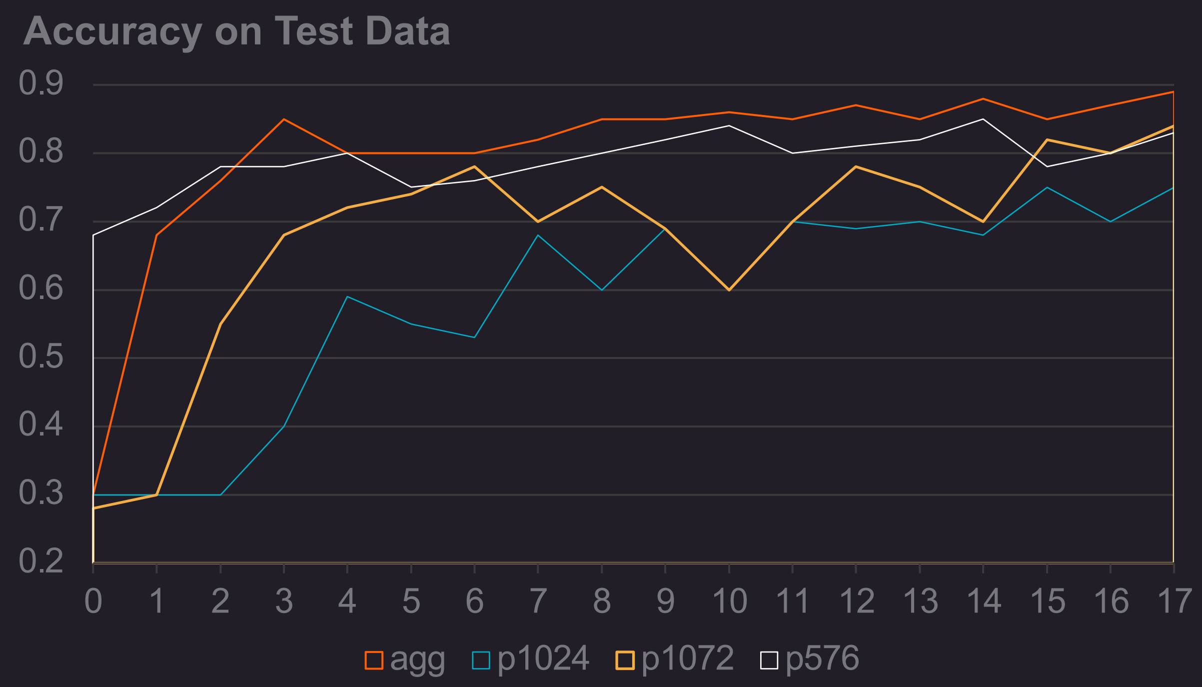
175
150
125
100
75
50
25
0

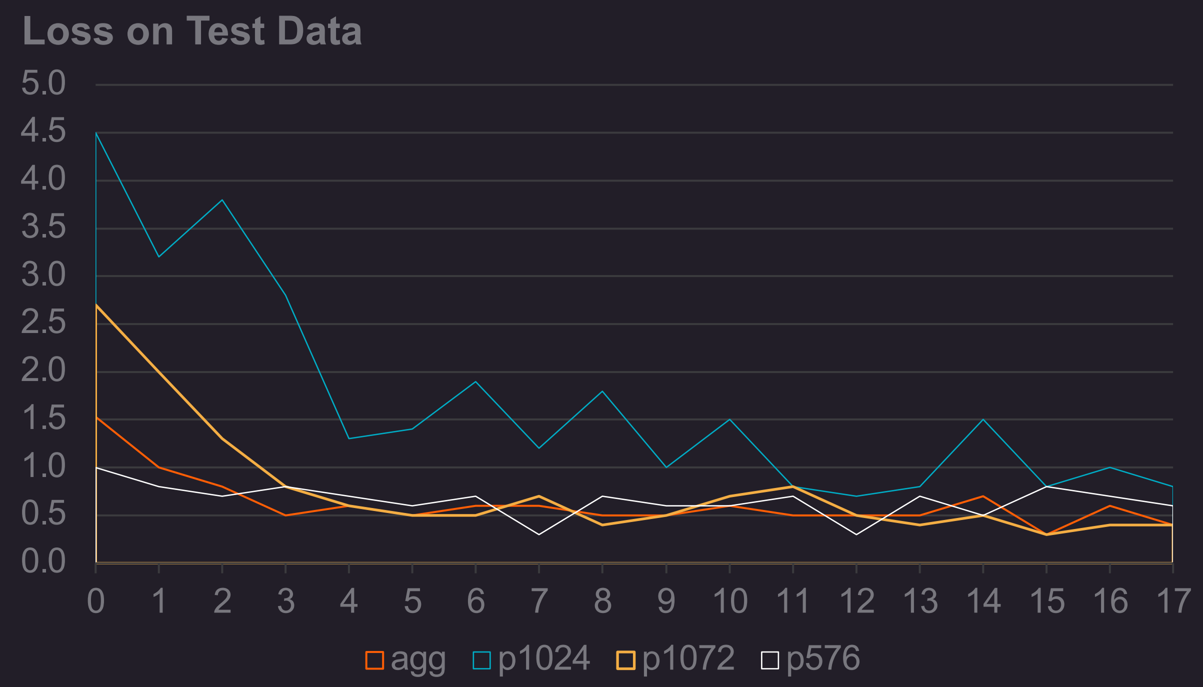
FL Model Performance on 1600 Images

Accuracy of Federated Model: 0.85 Despite Class Imbalance

Endoscopy Image Classification with Class Imbalance Results

Federated Learning Model





Accuracy and Loss vs Number of FL Rounds

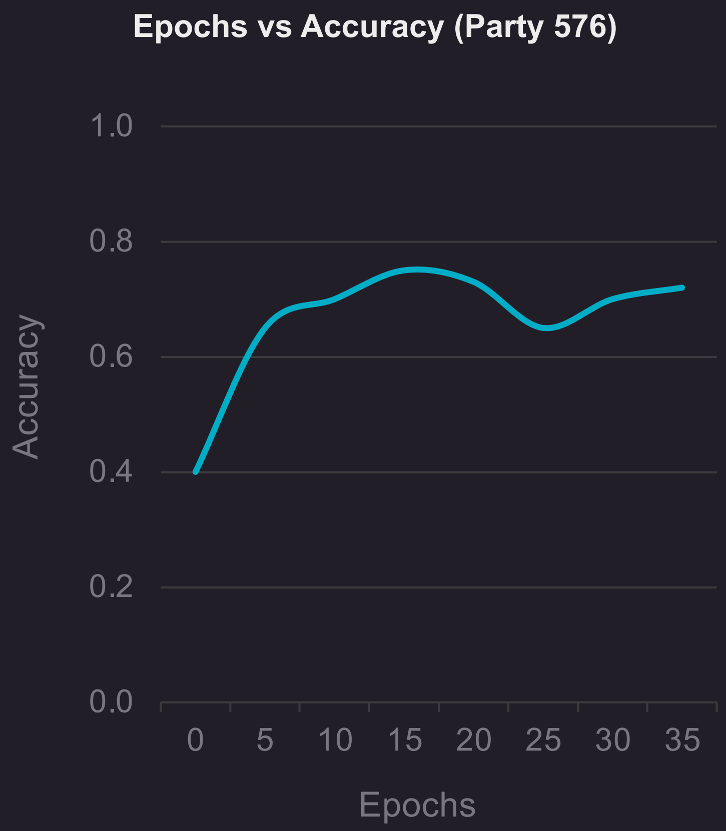
FL Model built with collaboration achieves Higher Accuracy

Test Data across all parties and Aggregator was kept same

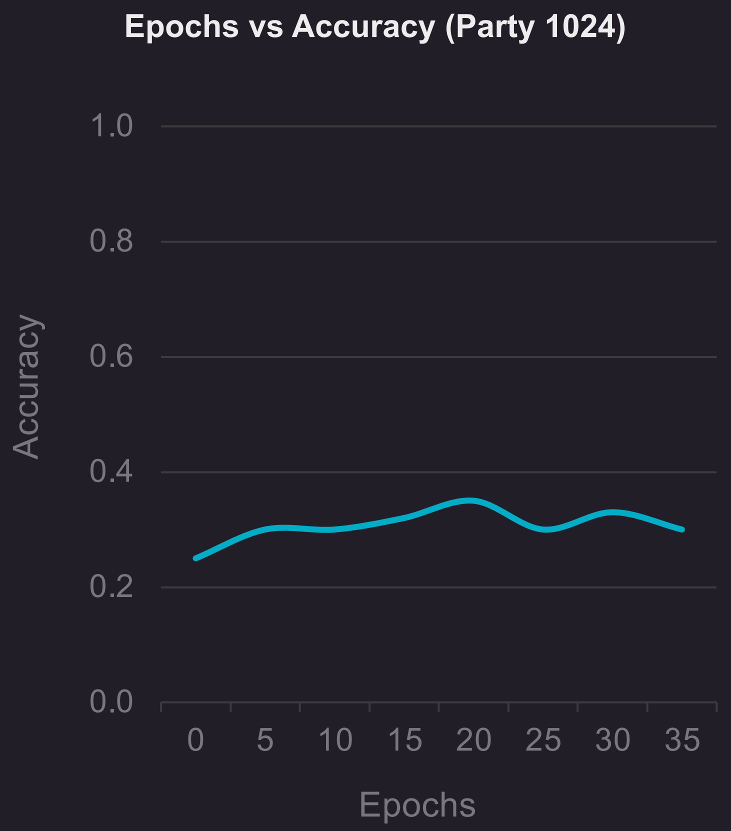
Endoscopy Image Classification with Class Imbalance Results (Contd.)

Non-FL Models

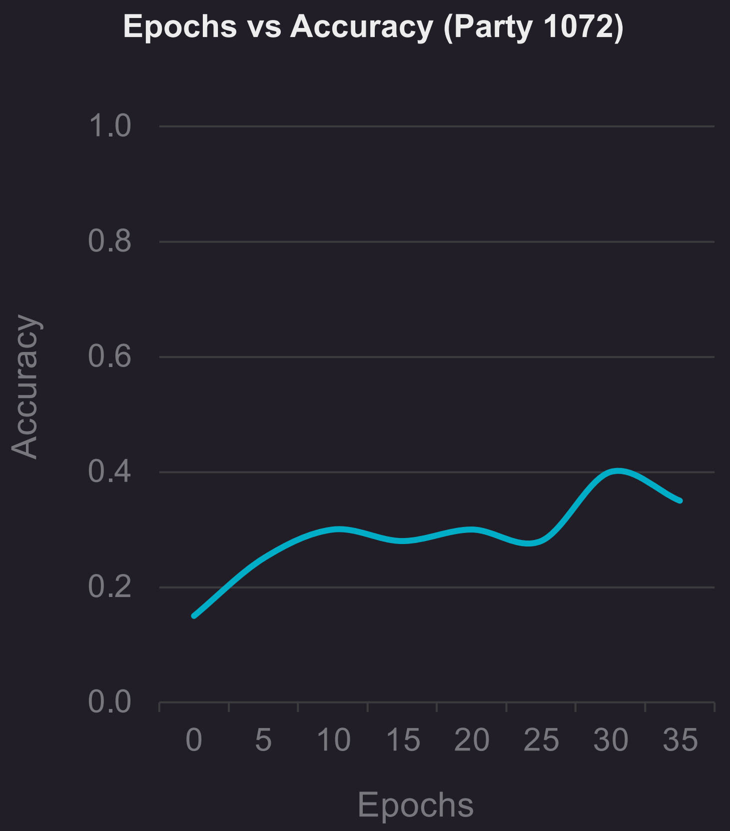
P576 Model



P1024 Model



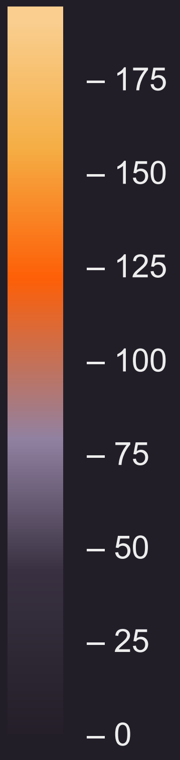
P1072 Model



Endoscopy Image Classification with Class Imbalance Results

Federated Learning Model Confusion Matrix

	0	1	2	3	4	5	6	7
0	175	0	0	0	1	2	0	0
1	0	194	0	0	1	0	19	0
2	16	1	183	2	6	8	0	1
3	0	0	0	174	0	0	0	38
4	0	0	2	2	190	4	0	4
5	19	0	4	1	7	173	1	0
6	4	31	1	0	0	5	147	0
7	0	0	2	48	0	3	0	131



- FL Model built with collaboration is least confused
- Test Data across all parties and Aggregator was kept same

Endoscopy Image Classification with Class Imbalance Results (Contd.)

Non-FL Models Confusion Matrix

P576 Model

	0	1	2	3	4	5	6	7
0	168	0	8	0	0	2	0	0
1	0	185	2	0	0	0	27	0
2	0	2	212	0	0	3	0	0
3	0	2	55	127	0	14	0	14
4	0	0	50	13	85	52	0	2
5	12	0	15	0	0	175	3	0
6	0	41	1	0	0	5	141	0
7	0	4	36	65	0	22	0	57

P1024 Model

	0	1	2	3	4	5	6	7
0	0	0	0	0	177	0	0	1
1	0	0	0	30	25	26	0	133
2	0	0	0	19	178	17	0	3
3	0	0	0	151	3	0	0	58
4	0	0	0	1	200	0	0	1
5	0	0	0	3	191	8	0	3
6	0	0	0	15	54	32	0	87
7	0	0	0	30	16	0	0	138

P1072 Model

	0	1	2	3	4	5	6	7
0	0	0	79	0	86	13	0	0
1	0	0	126	7	14	60	0	7
2	0	0	156	2	51	8	0	0
3	0	0	36	121	6	25	0	24
4	0	0	3	0	196	1	0	2
5	0	0	64	2	98	41	0	0
6	0	0	118	1	12	55	0	2
7	0	0	24	38	22	4	0	96

Colour scale: 200, 175, 150, 125, 100, 75, 50, 25, 0

Endoscopy Image Classification with Class Imbalance Results

Federated Learning Model Classification Report

Classification Report				
	Precision	Recall	F1-score	Support
0	0.82	0.98	0.89	178
1	0.86	0.91	0.88	214
2	0.95	0.84	0.89	217
3	0.77	0.82	0.79	212
4	0.93	0.94	0.93	202
5	0.89	0.84	0.87	205
6	0.88	0.78	0.83	188
7	0.75	0.71	0.73	184
Accuracy			**0.85**	**1600**
Macro Avg.	**0.86**	**0.85**	**0.85**	**1600**
Weighted Avg.	**0.86**	**0.85**	**0.85**	**1600**

- FL Model built with collaboration gives best accuracy of 85%. F1 scores have increased across the board as well.
- Individually non federated party specific models had following accuracy numbers:
  P576 was at 72%, P1024 at 31% and P1072 at 38%.

Endoscopy Image Classification with Class Imbalance Results (Contd.)

Non-FL Models Classification Report

P576 Model

Party 576

	Precision	Recall	F1-score	Support
0	0.93	0.94	0.94	178
1	0.79	0.86	0.83	214
2	0.56	0.98	0.71	217
3	0.62	0.60	0.61	212
4	1.00	0.42	0.59	202
5	0.64	0.85	0.73	205
6	0.82	0.75	0.79	188
7	0.78	0.31	0.44	184
Accuracy			**0.72**	**1600**
Macro Avg.	**0.77**	**0.71**	**0.70**	**1600**
Weighted Avg.	**0.76**	**0.72**	**0.70**	**1600**

P1024 Model

Party 1024

	Precision	Recall	F1-score	Support
0	0.00	0.00	0.00	178
1	0.00	0.00	0.00	214
2	0.00	0.00	0.00	217
3	0.61	0.71	0.66	212
4	0.24	0.99	0.38	202
5	0.10	0.04	0.06	205
6	0.00	0.00	0.00	188
7	0.33	0.75	0.45	184
Accuracy			**0.31**	**1600**
Macro Avg.	**0.16**	**0.31**	**0.19**	**1600**
Weighted Avg.	**0.16**	**0.31**	**0.19**	**1600**

P1072 Model

Party 1072

	Precision	Recall	F1-score	Support
0	0.00	0.00	0.00	178
1	0.00	0.00	0.00	214
2	0.26	0.72	0.38	217
3	0.71	0.57	0.63	212
4	0.40	0.97	0.57	202
5	0.20	0.20	0.20	205
6	0.00	0.00	0.00	188
7	0.73	0.52	0.61	184
Accuracy			**0.38**	**1600**
Macro Avg.	**0.29**	**0.37**	**0.30**	**1600**
Weighted Avg.	**0.29**	**0.38**	**0.30**	**1600**

Endoscopy Polyps Segmentation by Federating over Two Separate Datasets

Kvasir Polyps Images: ~1000, Norway Hospitals

CVC Polyps Images: ~600, Spain Hospitals

Improved segmentation model initially trained only with Spain Hospital Data by doing FL with Norway Hospital Data

Leveraged IBM FL for building models in a multi-party Federated Learning setting with FedAvg, Fed+ Fusion Algorithms

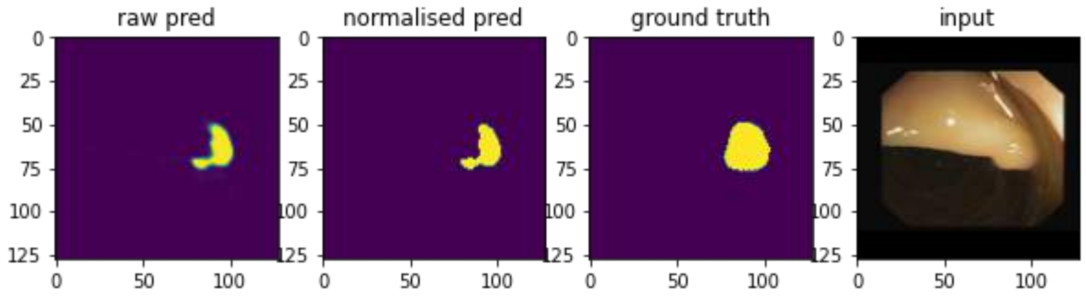
IBM

Federated Learning

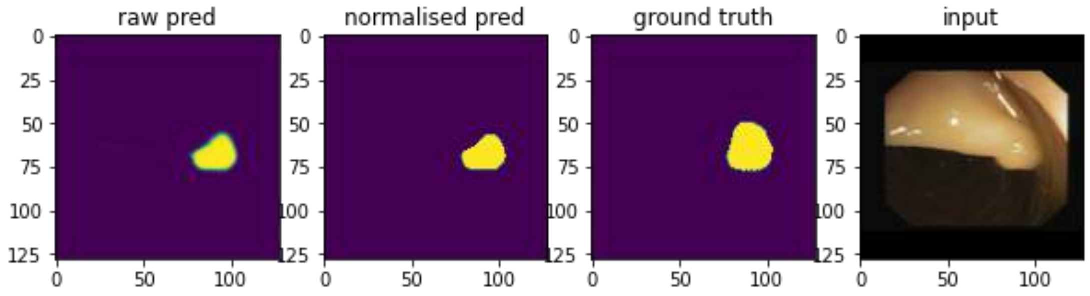
Source: CVC Clinic-DB

Persistent

Endoscopy Image Classification with Class Imbalance



Pre-Federated Learning Model
trained only on CVC Data



Post-Federated Learning Model
trained with insights from Kvasir Data

- Pre-Federated Learning, PSNR = 16.5
- Custom Loss Functions: PSNR

- FL Model Performance on 112 images from CVC Data
- PSNR increase from 16.5 to 18.4 , thanks to insights from Kvasir Data

Thumb Rules of Federated **Learning Applications**

Data is siloed and cannot be aggregated

Its too expensive to aggregate data (IoT Bandwidth)

One should be convinced that model will be improved by access to more diverse training data

References

Federated Learning

OpenMined

FedAI

Flower Federated Framework

NVIDIA CLARA

Google Comic Explanation

Predictive Maintenance

Credit Scoring

References

Cloudera Fast Forward Labs Federated Learning Report	**The Future of Digital Health with Federated Learning**	**IBM Federated Learning**
https://federated.fastforwardlabs.com	https://www.nature.com/articles/s41746-020-00323-1	https://ibmfl.mybluemix.net

Persistent

Thank you!