

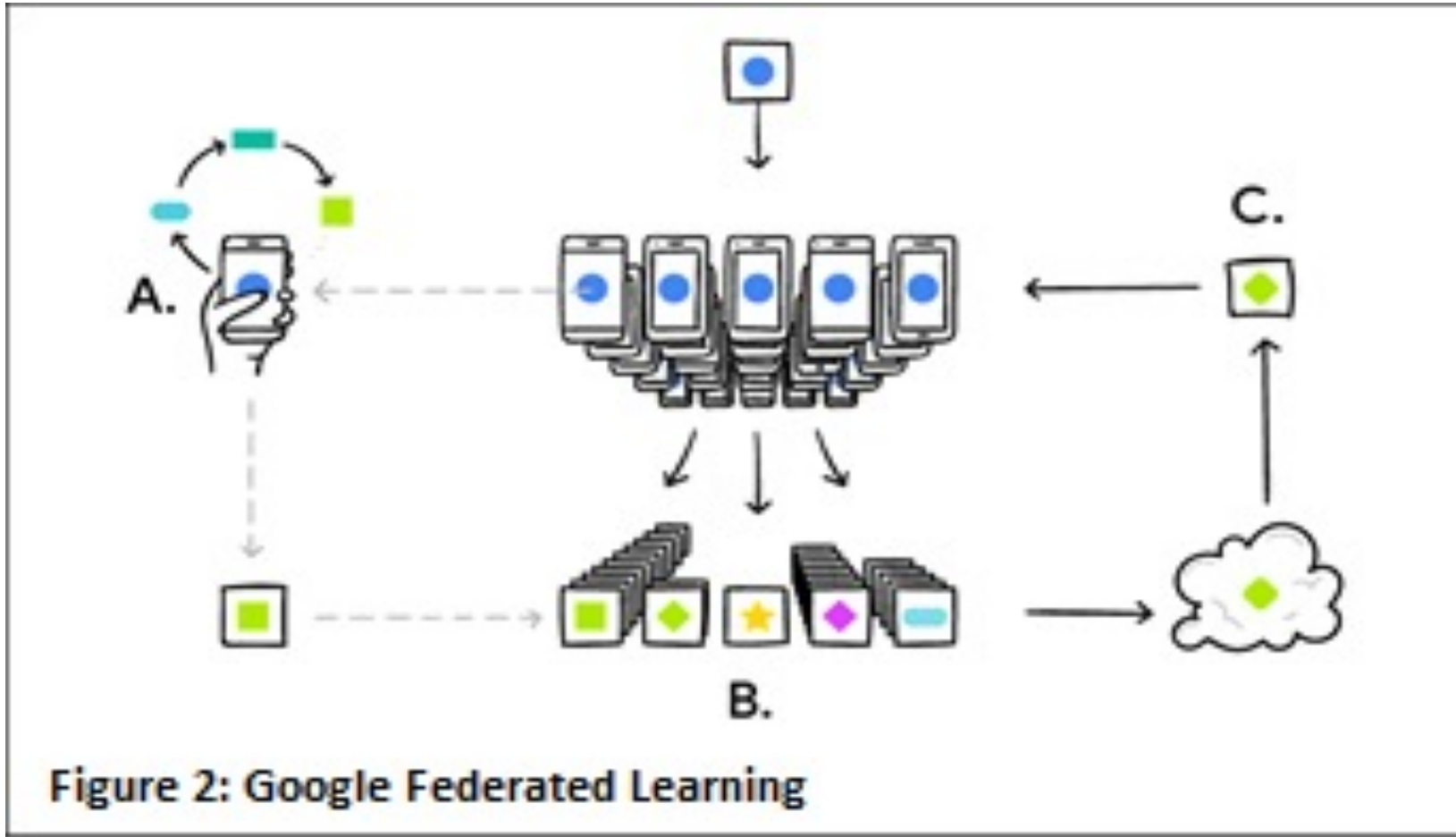


Institute
and Faculty
of Actuaries

Hush hush: Keeping neural network claims modelling private, secret, and distributed using federated learning

Dr Małgorzata Śmietanka & Dylan Liew

Smartphone Federated Learning Pipeline



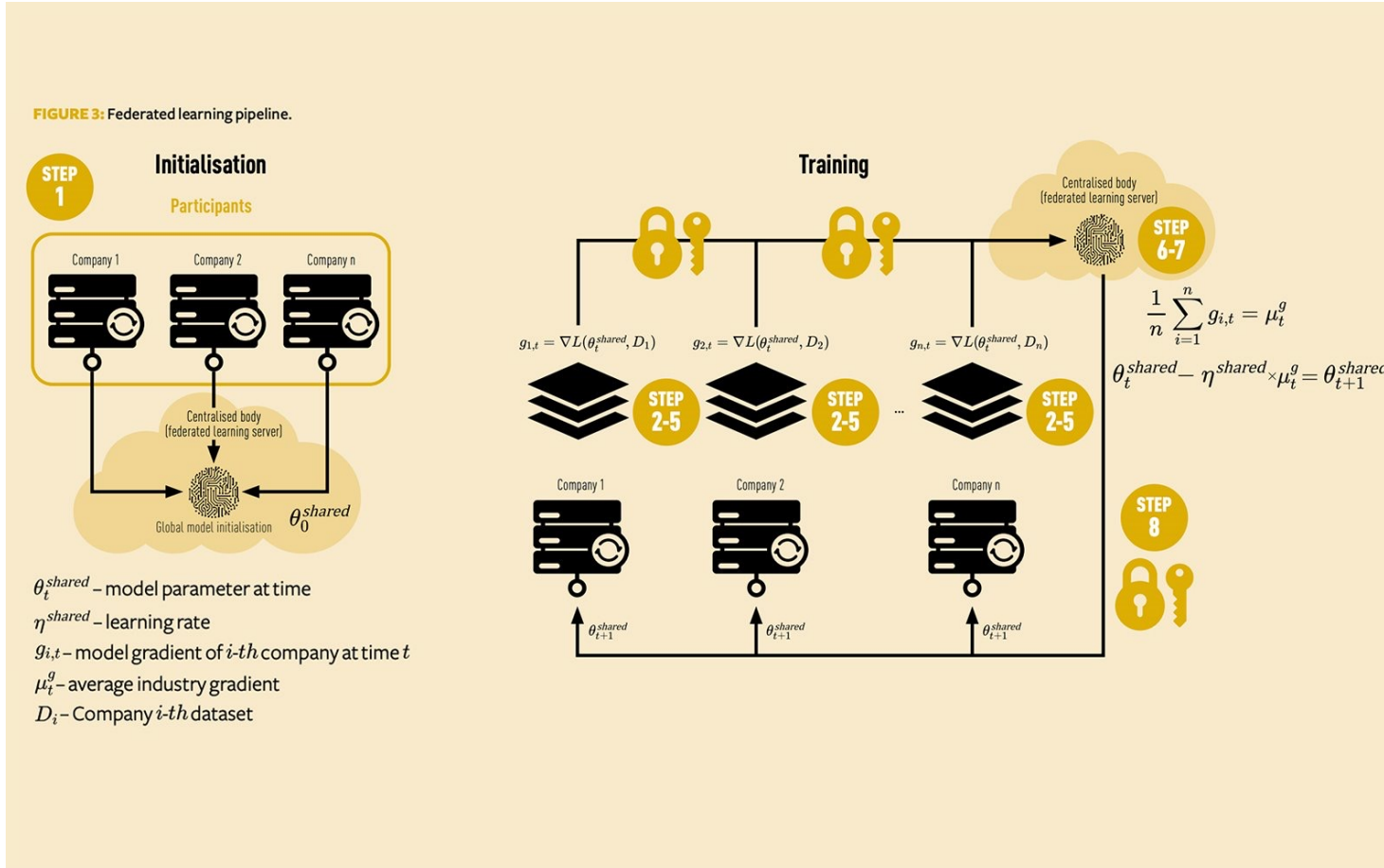
A) your phone personalises the model locally depending on your usage;
B) many users' updates are aggregated;
C) the aggregated updates form a consensus change to the shared model; and
D) the shared models are updated.

[Google AI Blog: Federated Learning: Collaborative Machine Learning without Centralized Training Data \(googleblog.com\)](https://googleblog.com)

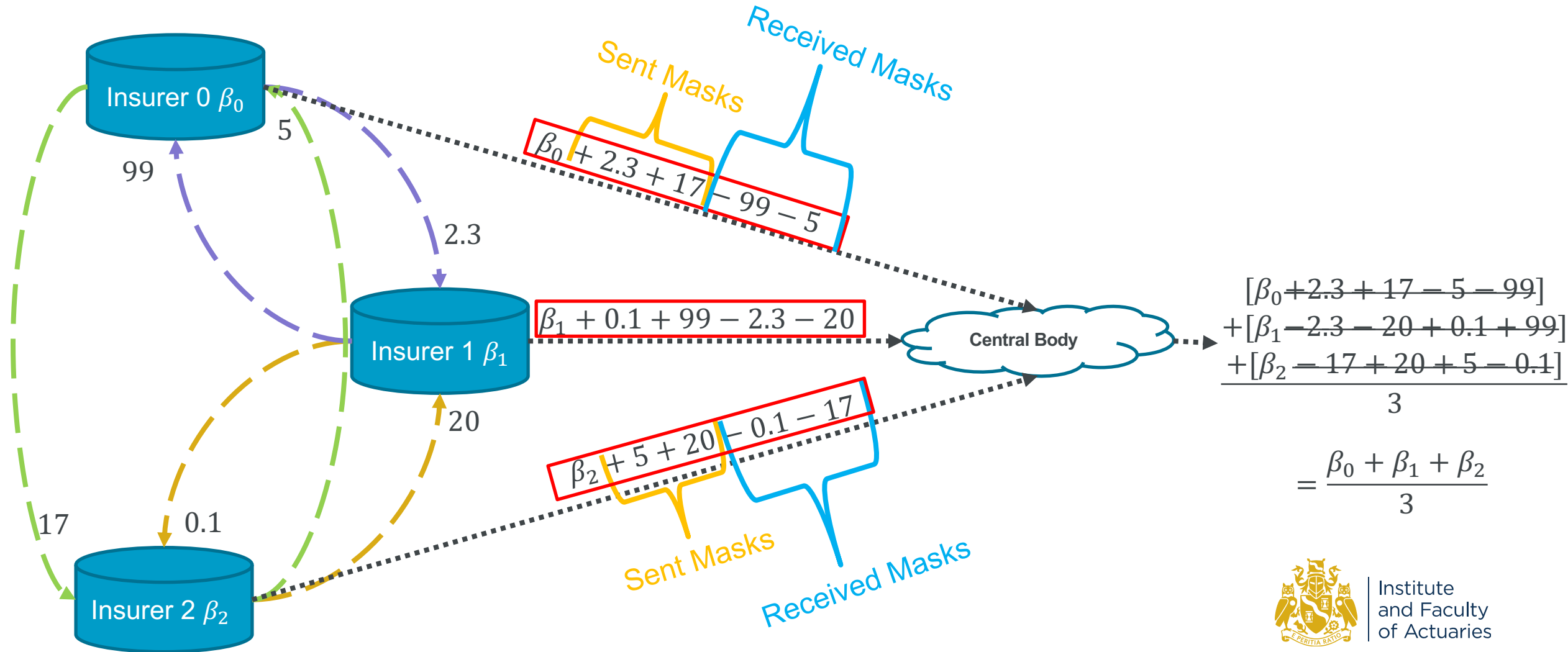


Institute
and Faculty
of Actuaries

Insurance Federated Learning Pipeline



Need to encrypt parameters but maintain the average

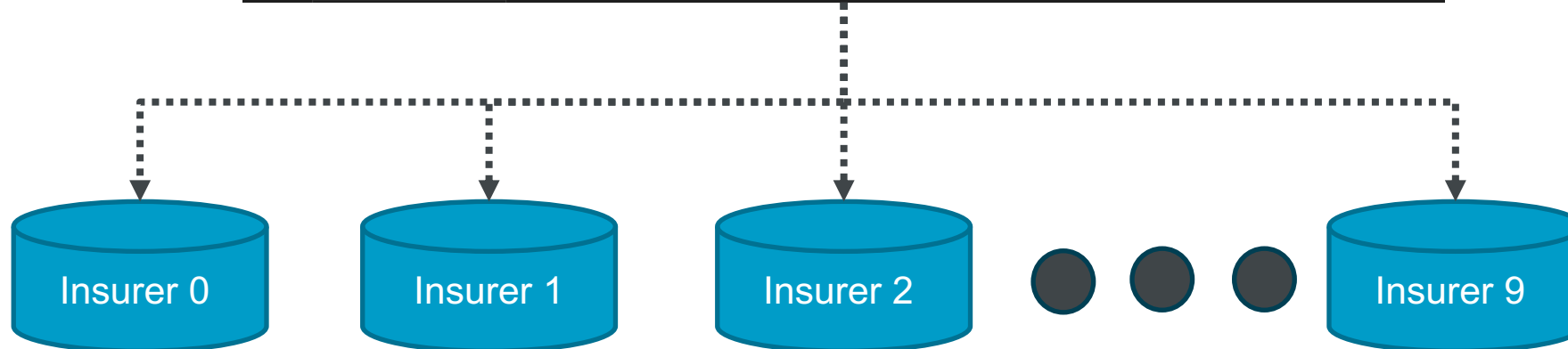


Insurance Federated Learning Pipeline

```
df = pd.read_csv('../data/FreMTP2Freq.CSV')
df
[108] ✓ 0.5s
```

| | IDpol | ClaimNb | Exposure | Area | VehPower | VehAge | DrivAge | BonusMalus | VehBrand | VehGas | Density | Region |
|--------|-----------|---------|----------|------|----------|--------|---------|------------|----------|---------|---------|--------|
| 0 | 1.0 | 1 | 0.10000 | D | 5 | 0 | 55 | 50 | B12 | Regular | 1217 | R82 |
| 1 | 3.0 | 1 | 0.77000 | D | 5 | 0 | 55 | 50 | B12 | Regular | 1217 | R82 |
| 2 | 5.0 | 1 | 0.75000 | B | 6 | 2 | 52 | 50 | B12 | Diesel | 54 | R22 |
| 3 | 10.0 | 1 | 0.09000 | B | 7 | 0 | 46 | 50 | B12 | Diesel | 76 | R72 |
| 4 | 11.0 | 1 | 0.84000 | B | 7 | 0 | 46 | 50 | B12 | Diesel | 76 | R72 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 678008 | 6114326.0 | 0 | 0.00274 | E | 4 | 0 | 54 | 50 | B12 | Regular | 3317 | R93 |
| 678009 | 6114327.0 | 0 | 0.00274 | E | 4 | 0 | 41 | 95 | B12 | Regular | 9850 | R11 |
| 678010 | 6114328.0 | 0 | 0.00274 | D | 6 | 2 | 45 | 50 | B12 | Diesel | 1323 | R82 |
| 678011 | 6114329.0 | 0 | 0.00274 | B | 4 | 0 | 60 | 50 | B12 | Regular | 95 | R26 |
| 678012 | 6114330.0 | 0 | 0.00274 | B | 7 | 6 | 29 | 54 | B12 | Diesel | 65 | R72 |

678013 rows x 12 columns



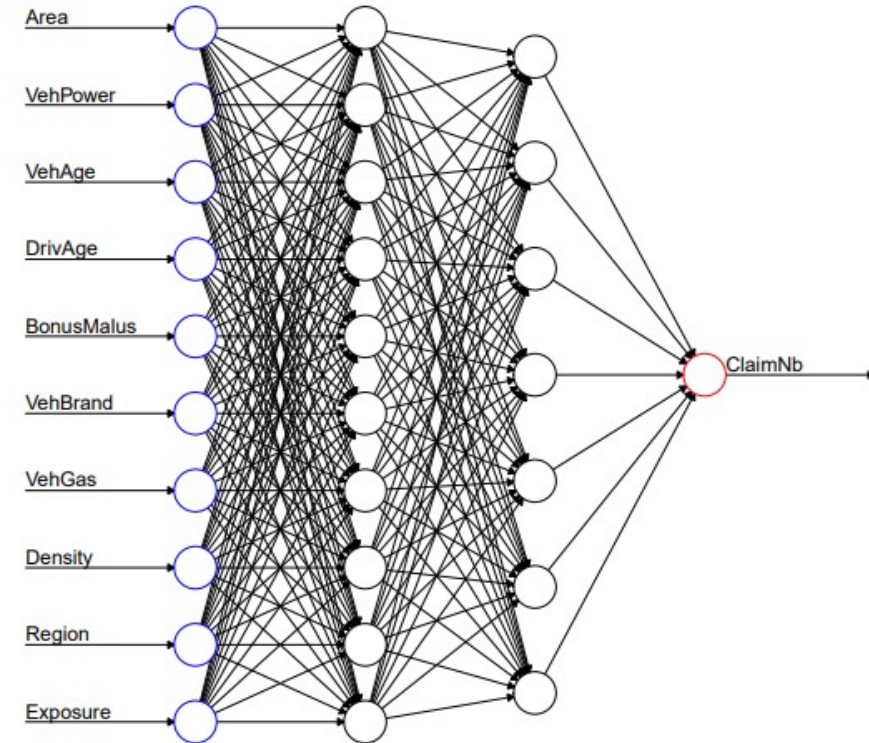
Neural Network Model Setup

Table 2. Neural Network Architecture used in all 3 Scenarios

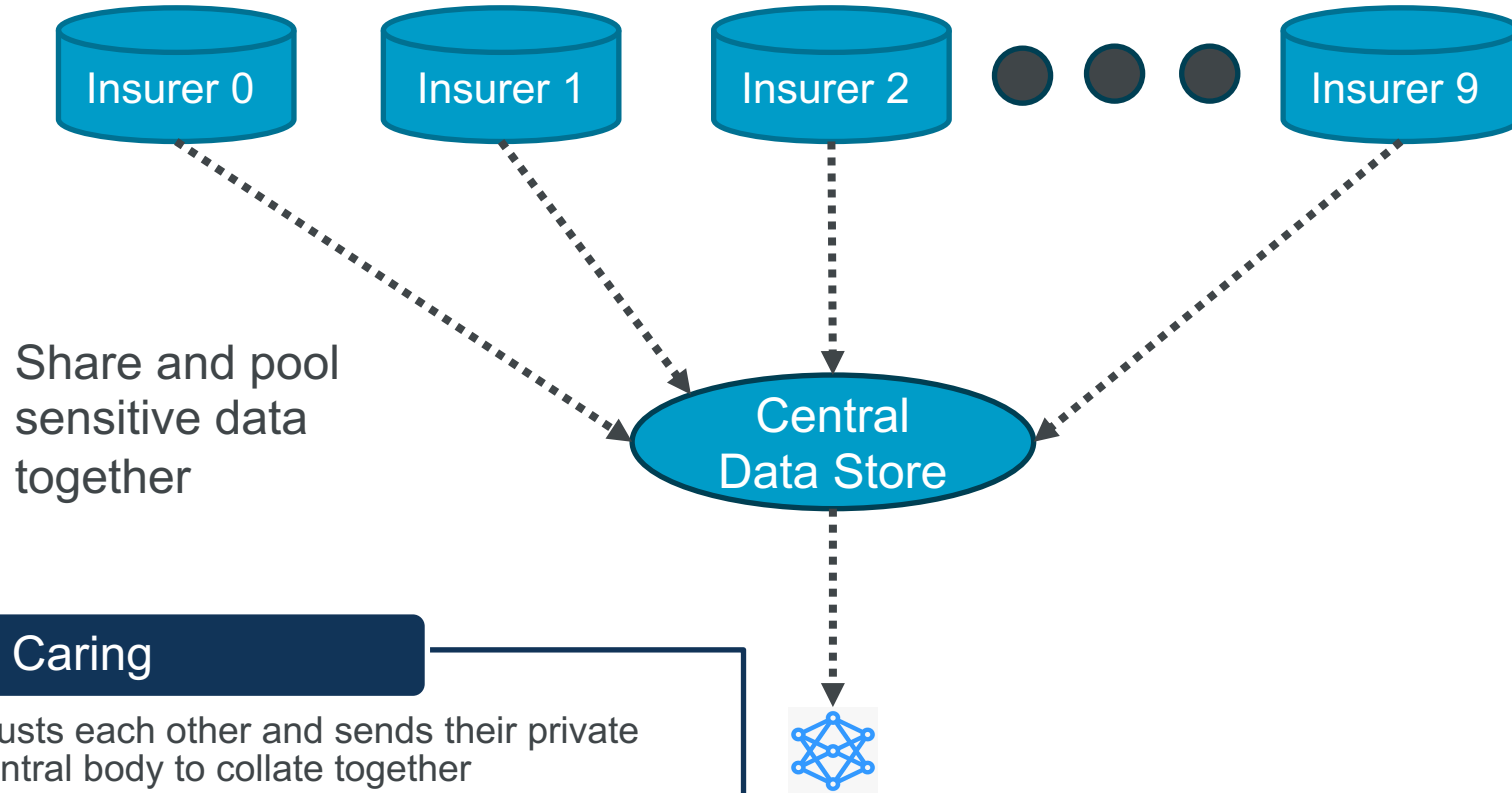
| Hyperparameter | Selection |
|---------------------|--|
| Input neurons | 39 based on the preprocessing done in Section 5.2.2 |
| Hidden Layers | 2 |
| Output Layer | 1 output neuron with exponential link function (to ensure only positive frequencies are predicted) |
| Optimiser | <i>NAdam</i> |
| Activation Function | <i>tanh</i> |
| Loss Function | Negative Poisson Log Likelihood |
| Initialisation | Xavier |
| Epochs | 300 |

Table 3. Hyperparameter Search Space Considered in all 3 Scenarios

| Hyperparameter | Search Space |
|----------------------------------|-----------------------------|
| Learning Rate | [0.001, 0.002, 0.01] |
| Number neurons in Hidden Layer 1 | [5, 10, 15, 20] |
| Number neurons in Hidden Layer 2 | [5, 10, 15, 20] |
| Batch Size | [500, 1,000, 5,000, 10,000] |



Global Model Scenario – 10 insurers, 1 models

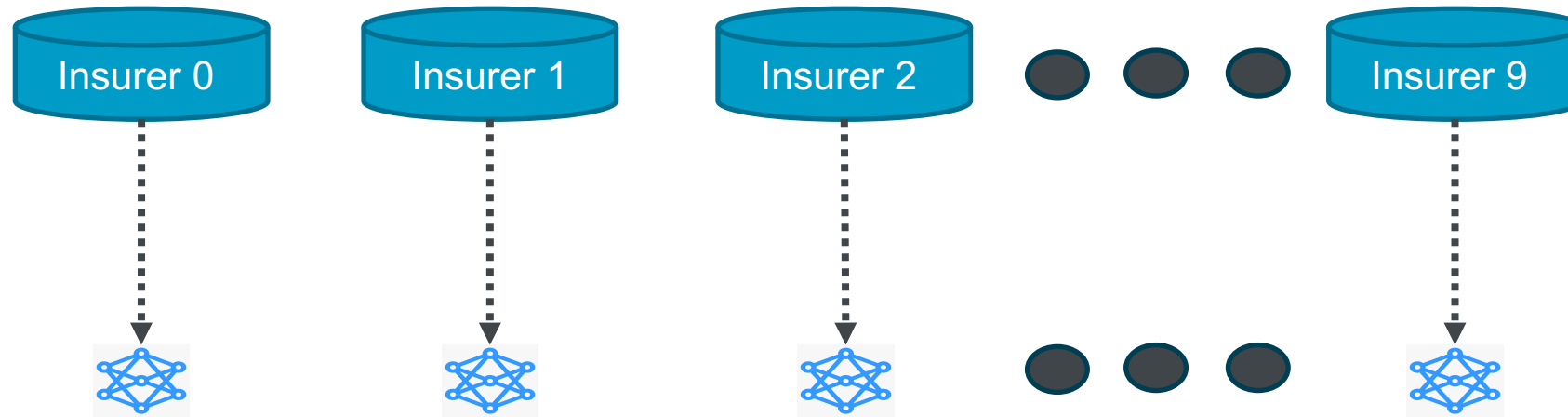


Sharing is Caring

- Everyone trusts each other and sends their private data to a central body to collate together
- Central body builds model for everyone and then sends back to companies
- **A.k.a. 1 “Global” model** as it uses all the data and applies to everyone



Partial Model Scenario – 10 insurers, 10 models

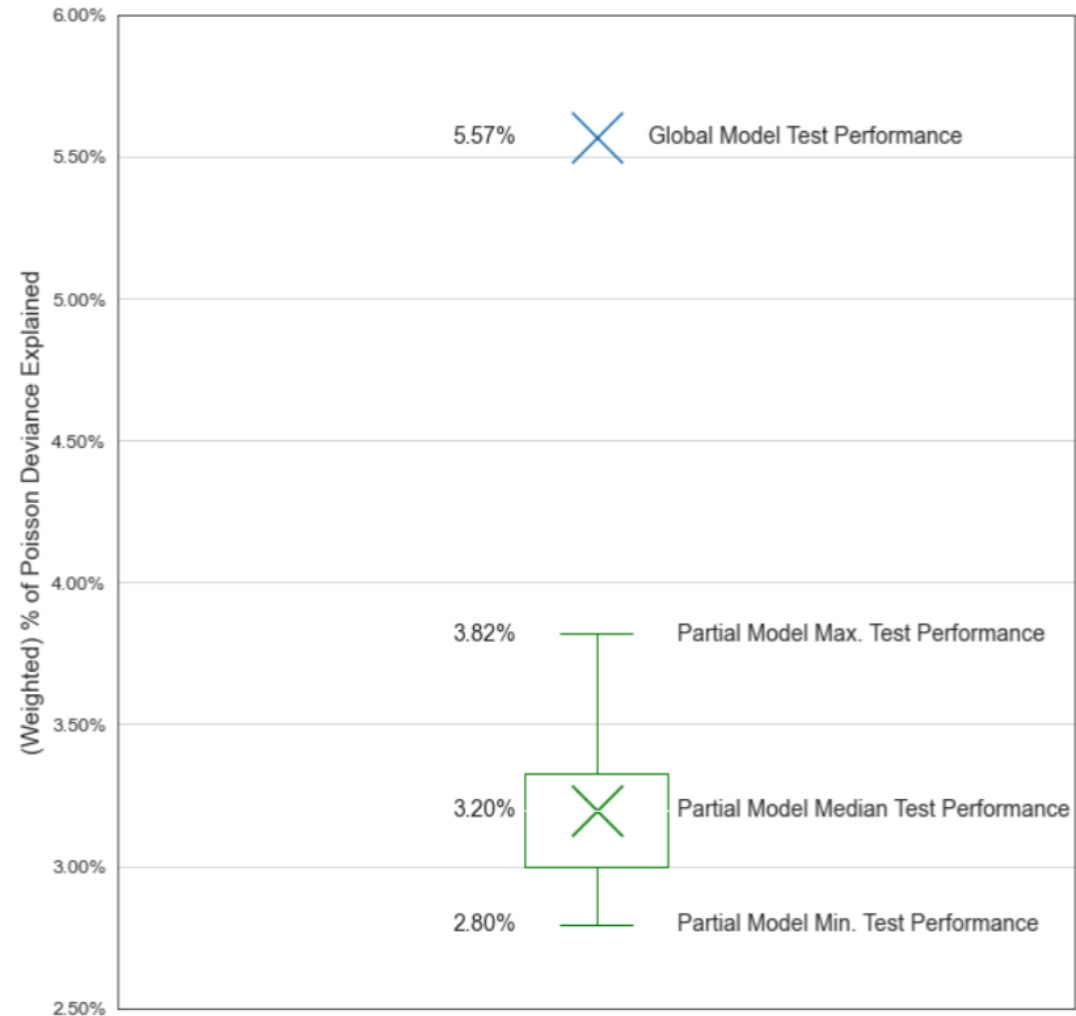


Each insurer builds their own model just using their data

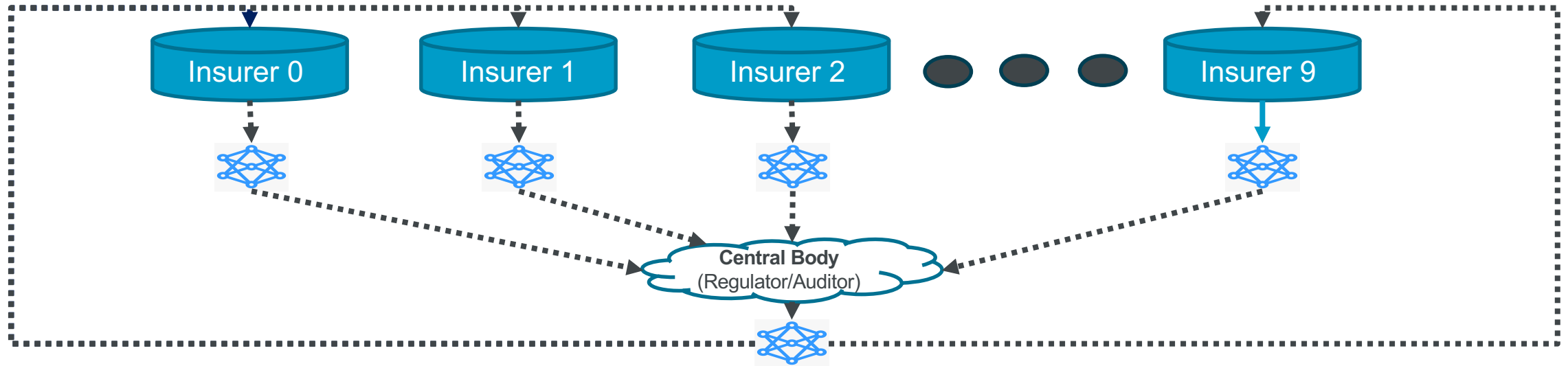
- No one trusts anyone
- Low volume of data used to build models which could be more relevant to company although may not be credible
- **A.k.a. 10 “Partial” models** as each company's model only has partial access to the whole market data



If you could share data



Federated Model Scenario – 10 insurers, 1 model

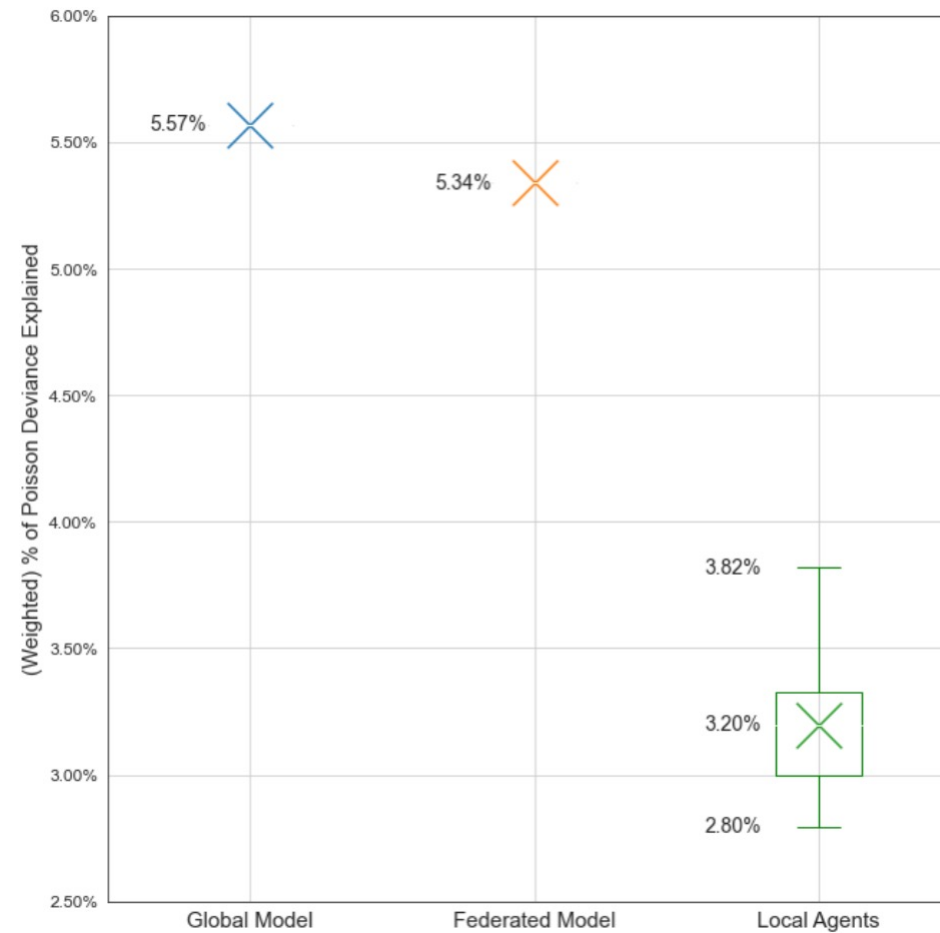


“United Federation”

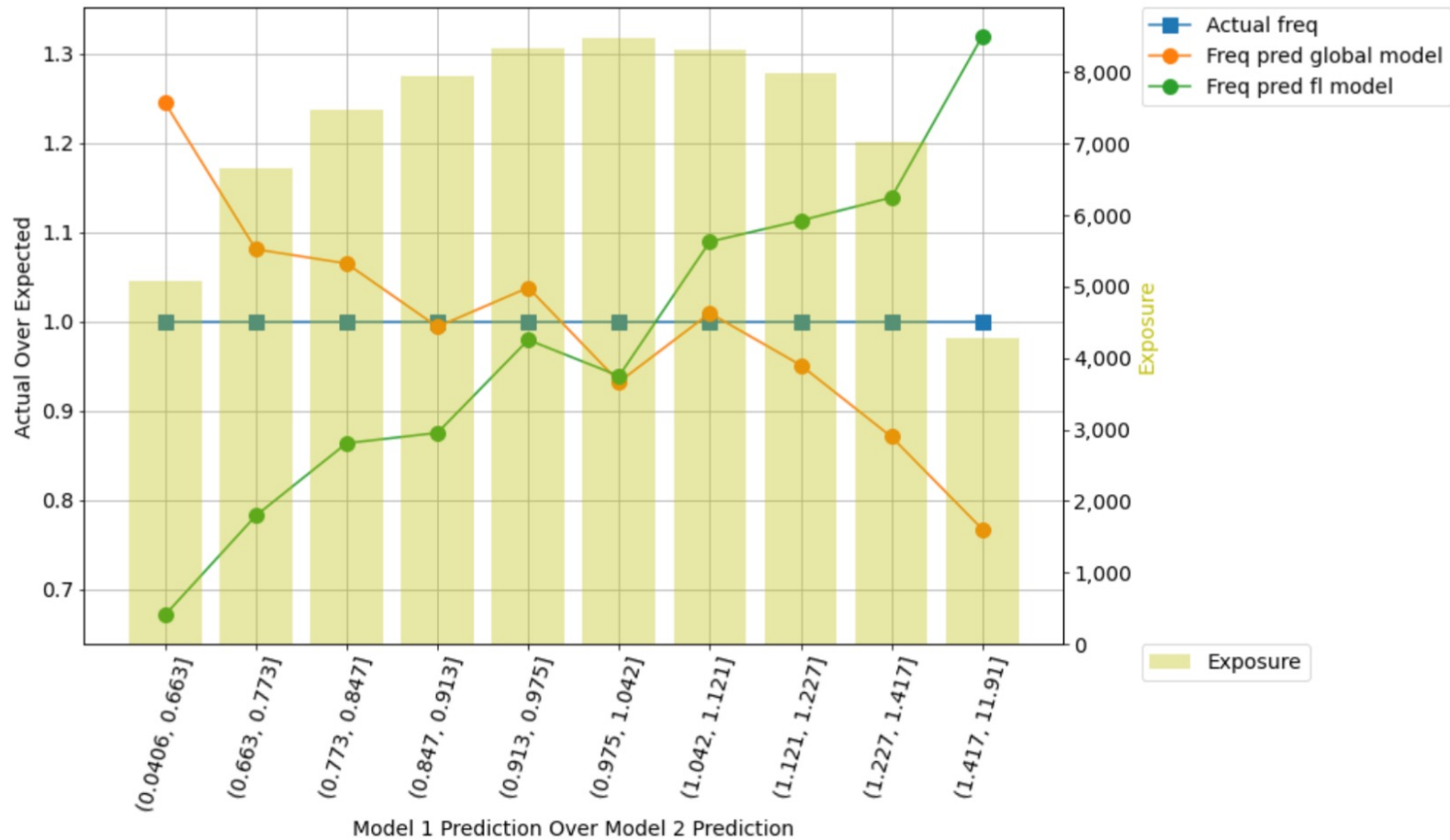
- Everyone keeps their 10th of their data to themselves
- However they securely share their parameters with central body
- Central body securely averages all the insurer's parameters and shares back
- Bringing the model to the data rather than bringing the data to the model
- **A.k.a. 1 “Federated” model**



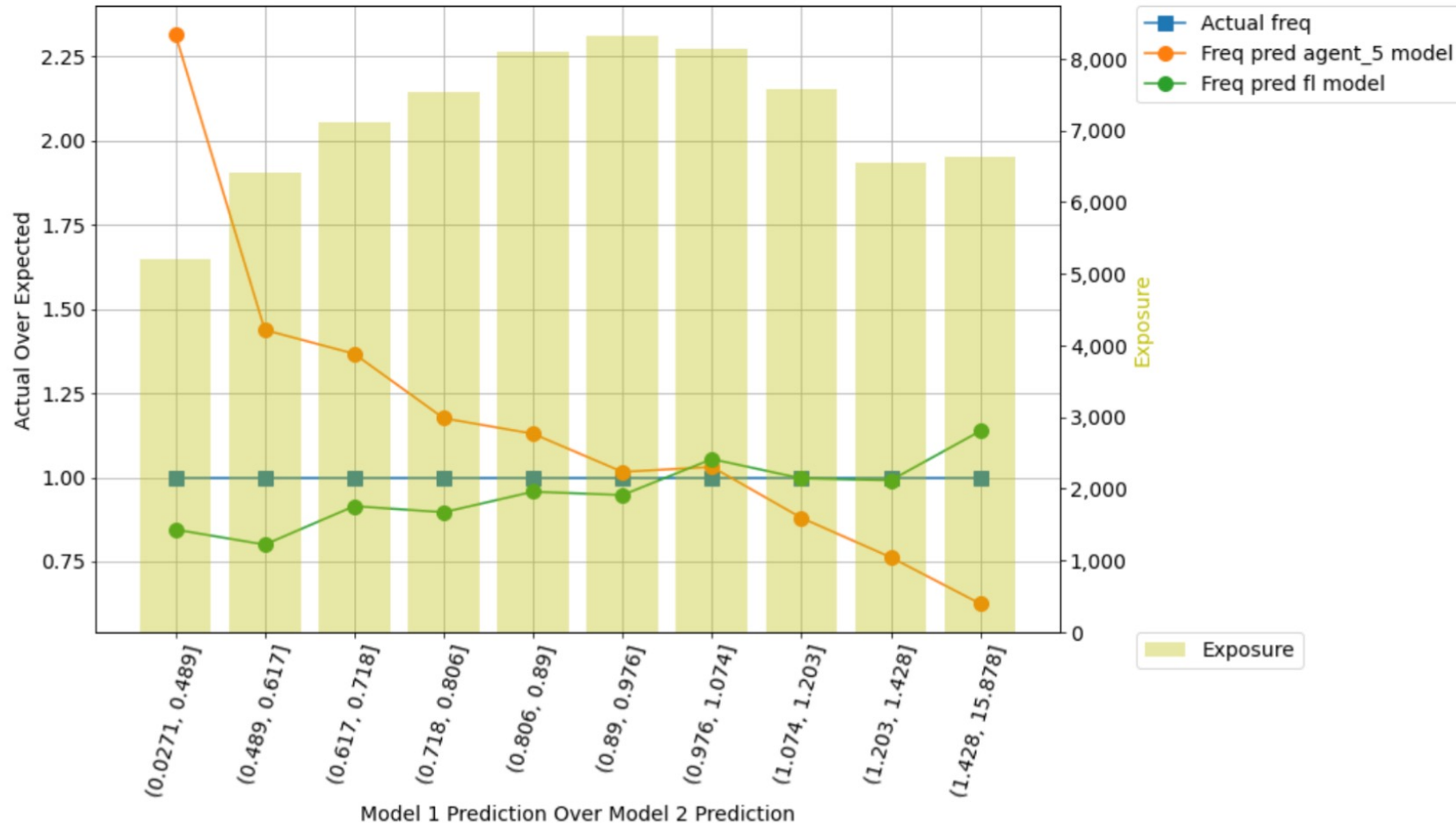
Comparison of results



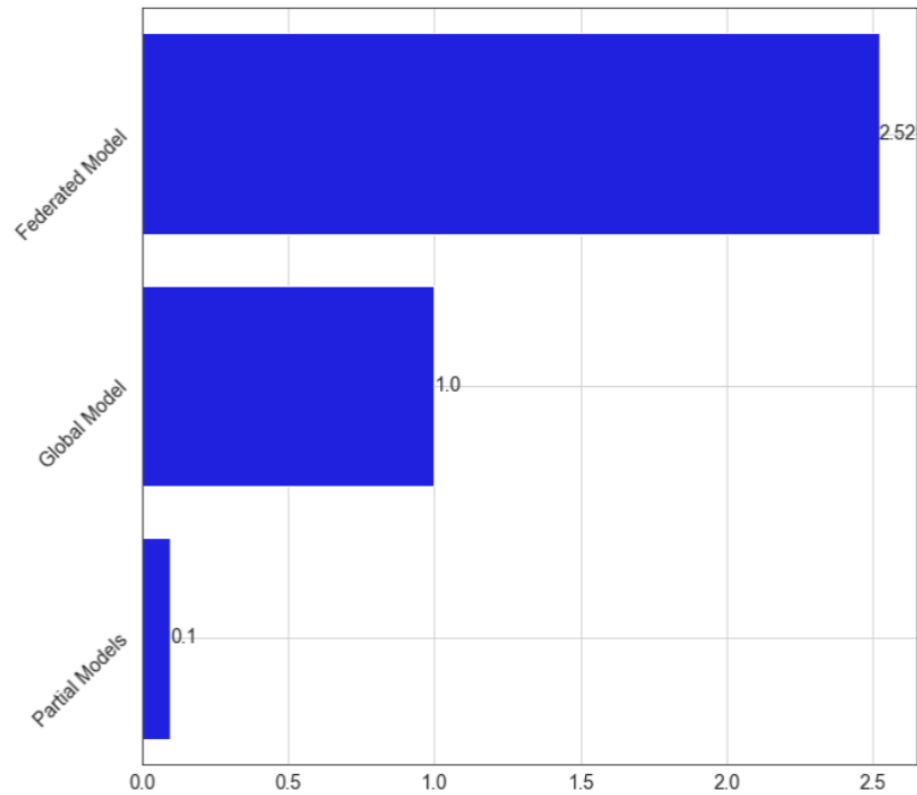
Comparison of results



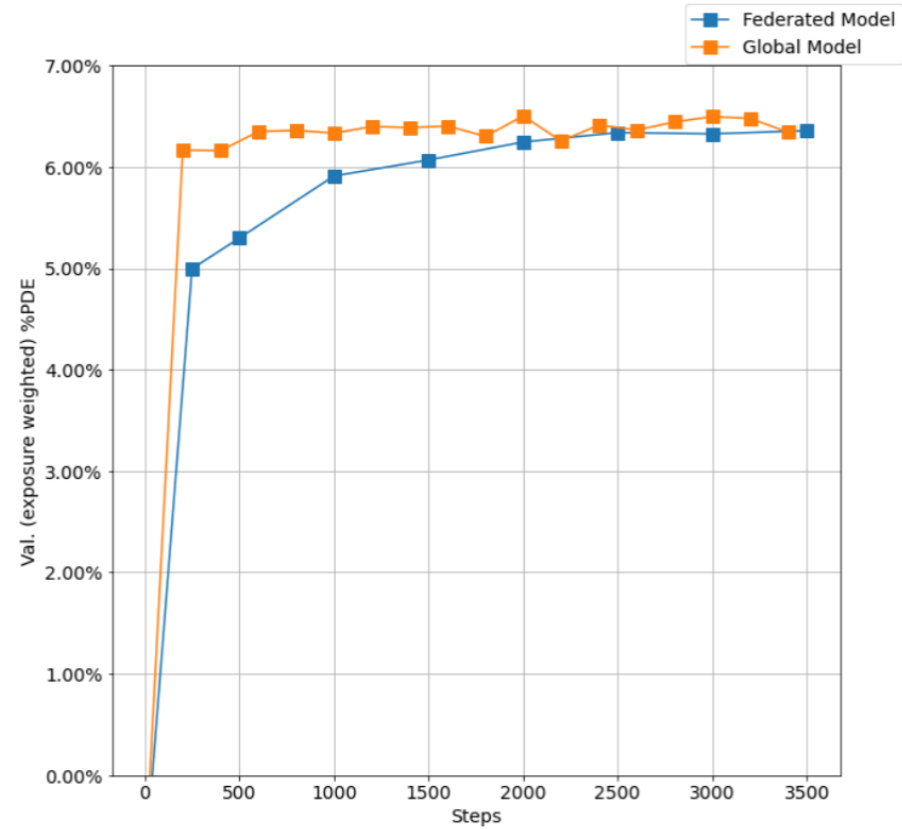
Comparison of results



What's the catch?



(a) Relative increase in observed wall time to train the models compared to training the Global Model.



(b) Exposure Weighted Validation % PDE of the Global and Federated Models over different number of parameter update steps.



Questions

Comments

Expressions of individual views by members of the Institute and Faculty of Actuaries and its staff are encouraged.

The views expressed in this presentation are those of the presenter.



Institute
and Faculty
of Actuaries