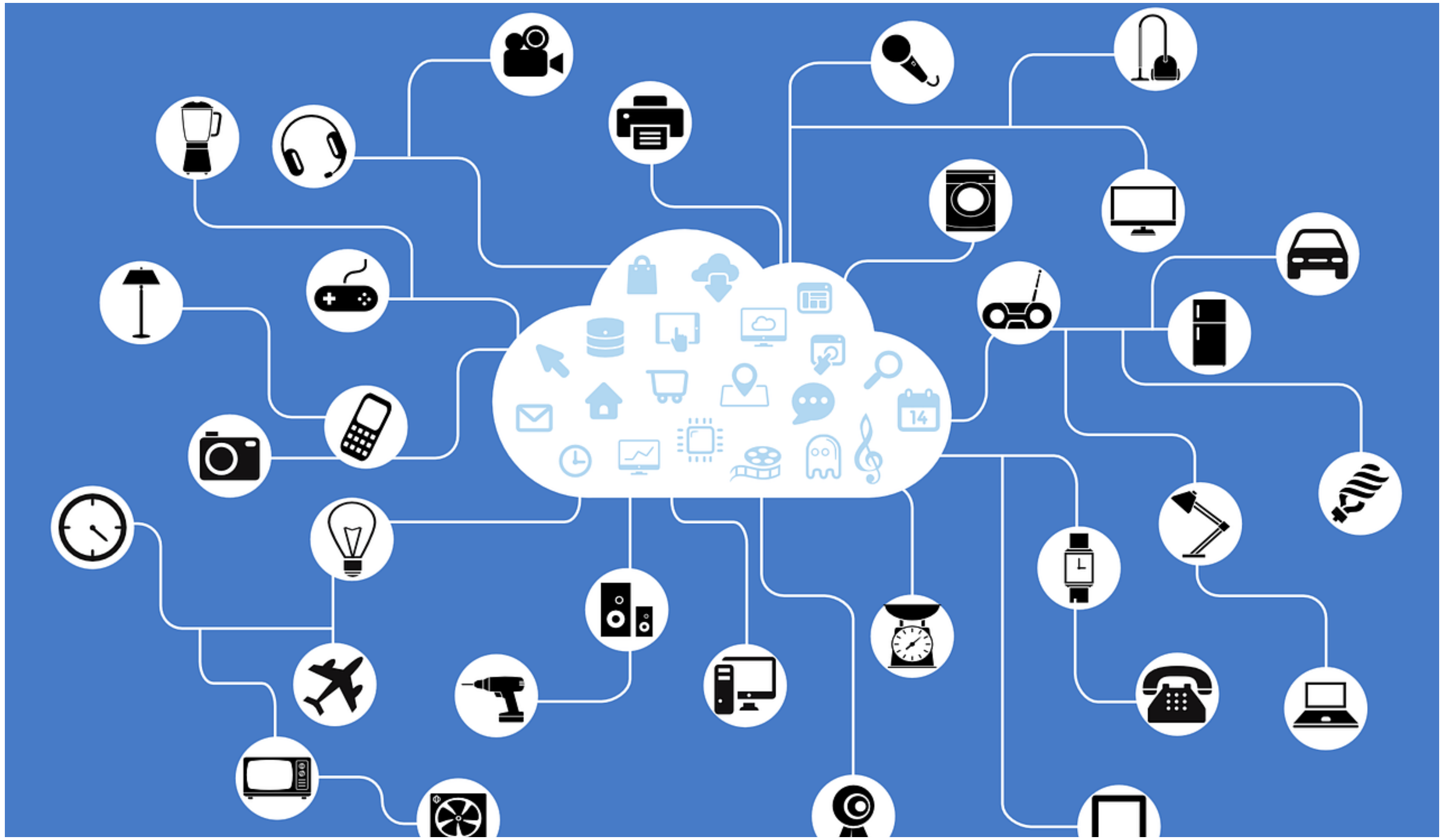
RESOURCE-CONSTRAINED FEDERATED LEARNING WITH HETEROGENEOUS LABELS AND MODELS

Gautham Krishna Gudur
Ericsson R&D

Bala Shyamala Balaji
Indian Institute of Technology, Madras

Perepu Satheesh Kumar
Ericsson Research



IoT ON THE EDGE

- Expansive growth of usage IoT devices with multiple sensors across various users.
- Significant research in the field of edge and ubiquitous computing.
- Data from sensors conveniently provide a way to extract contextual, behavioural information of users.

LEARNING FROM MULTIPLE DEVICES ON THE EDGE

Collaborative and Distributed Machine Learning is now possible more than ever to help best utilize the information learnt from multiple IoT devices.

Practical Challenges

Privacy Concerns about sharing sensitive data to the cloud from local user devices	***Low Latency*** between cloud and local devices

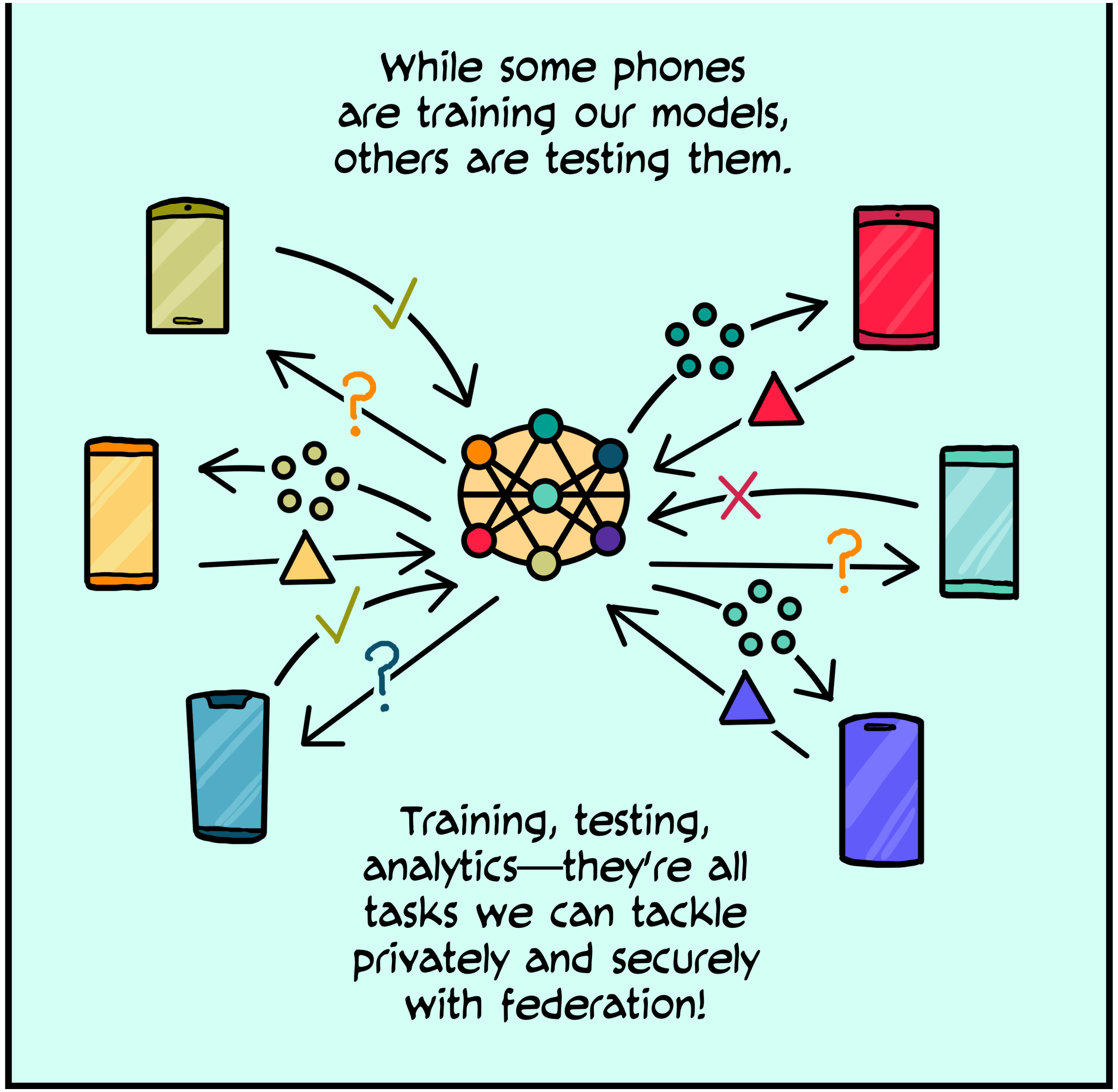
FEDERATED LEARNING

Algorithms are trained across a federation of multiple decentralized devices.

Effectively train a global/centralized model without compromising on sensitive data of various users.

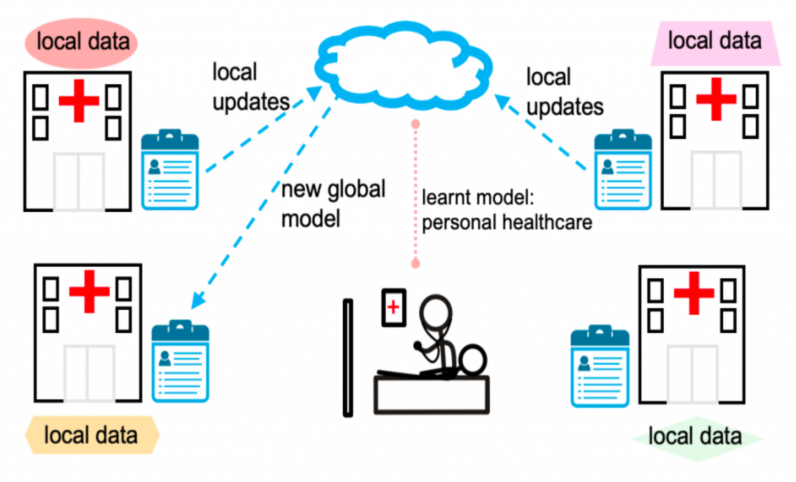
Transfer of model weights and updates from local devices to cloud, rather than conventional sharing of data.

More Personalization; Minimal Latency; Privacy Preserving.



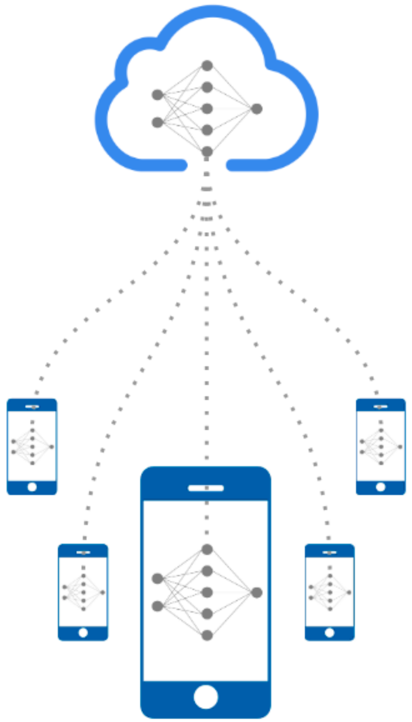
Picture taken from federated.withgoogle.com

HEALTHCARE

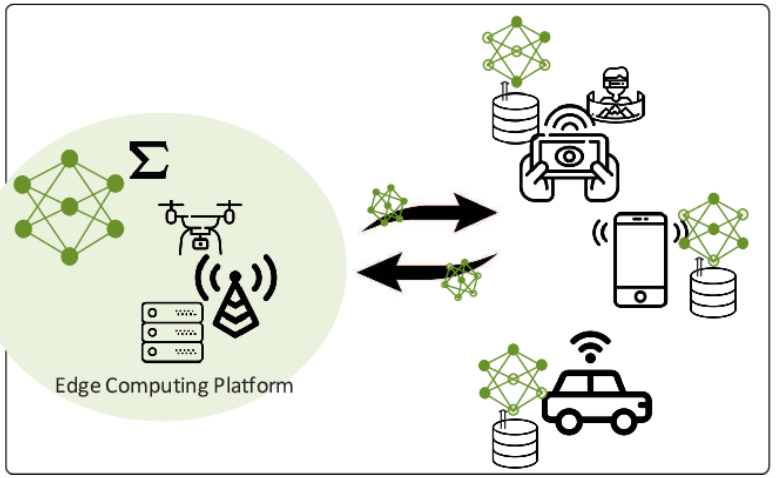


Picture taken from https://blog.ml.cmu.edu/2019/11/12/federated-learning-challenges-methods-and-future-directions/

IoT ON THE EDGE



WIRELESS/TELECOM



Picture taken from https://arxiv.org/pdf/1908.06847.pdf

FEDERATED LEARNING APPLICATIONS

PROMINENT CHALLENGES IN RESOURCE-CONSTRAINED FEDERATED LEARNING

- Communication Overheads – Reducing Latency
- Privacy Concerns – Sensitive Data Transfer
- Systems Heterogeneities – Hardware, Software, Network, Power (Resource Constraints)
- Statistical Heterogeneities
  - Non-IIDness
  - Model Heterogeneities

PROMINENT CHALLENGES IN RESOURCE-CONSTRAINED FEDERATED LEARNING

- Communication Overheads – Reducing Latency
- Privacy Concerns – Sensitive Data Transfer
- Systems Heterogeneities – Hardware, Software, Network, Power (Resource Constraints)
- Statistical Heterogeneities
  - Non-IIDness
  - Model Heterogeneities
  - ***Label Heterogeneities***

What are *Label Heterogeneities?*

PROMINENT CHALLENGES IN RESOURCE-CONSTRAINED FEDERATED LEARNING

- Communication Overheads – Reducing Latency
- Privacy Concerns – Sensitive Data Transfer
- Systems Heterogeneities – Hardware, Software, Network, Power (Resource Constraints)
- Statistical Heterogeneities
  - Non-IIDness
  - Model Heterogeneities
  - ***Label Heterogeneities***

What are *Label Heterogeneities*?

The flexibility to handle different labels across user devices.

GOALS OF OUR PROPOSED SYSTEM

- A framework to allow ***flexible heterogeneous selection of labels***, thereby leveraging information pertaining to specific classes (with and without label overlap).
- Flexibility in ***preferred local model architectures*** in a federated learning setting, for effective transfer learning between global and local models.
- Empirical demonstration of the framework's ability to handle different data distributions (***statistical heterogeneities and non-IID***) across various user devices.
- Demonstrating the ***feasibility of on-device personalized federated learning***, and resource-friendly; independent of users (*User Adaptability*).

PROPOSED FRAMEWORK

- *Model scores*, instead of model weights are sent to the cloud during every federated learning iteration.
- ***Build:*** We build the model on the incoming data pertaining to each local user at specific iteration.
- ***Local Update:*** Weighted average of scores across different iterations on same user.
  - ➢ We propose a weighted α-update, where α is the ratio between the size of current private dataset and the size of public dataset.
  - ➢ Governs the contributions of the new and old models.
- ***Global Update:*** Weighted average of scores across all users in same iteration.
  - ➢ β parameter governs the weightage given to overlapping labels across users.

Algorithm 1: Proposed Framework to handle heterogeneous labels and models in Federated Learning

---

Input - Public Data set $\mathcal{D}_0\{x_0, y_0\}$, Private datasets $\mathcal{D}_m^i$, Total users M, Total iterations I, LabelSet for each user l_m

Output - Trained Model scores f_G^I

Initialize - $f_G^0 = \mathbf{0}$ (Global Model Scores)

for $i = 1$ **to** I **do**

  for $m = 1$ **to** M **do**

    Build: Model $\mathcal{D}_m^i$ and predict $f_{\mathcal{D}_m^i}(x_0)$

    Local Update: $f_{\mathcal{D}_m^i}(x_0) = f_G^I(x_0^{l_m}) + \alpha f_{\mathcal{D}_m^i}(x_0)$, where $f_G^I(x_0^{l_m})$ are the global scores of only the set of labels l_m with the m^{th} user, and $\alpha = \frac{len(\mathcal{D}_m^i)}{len(\mathcal{D}_0)}$

    Global Update: Update label wise,

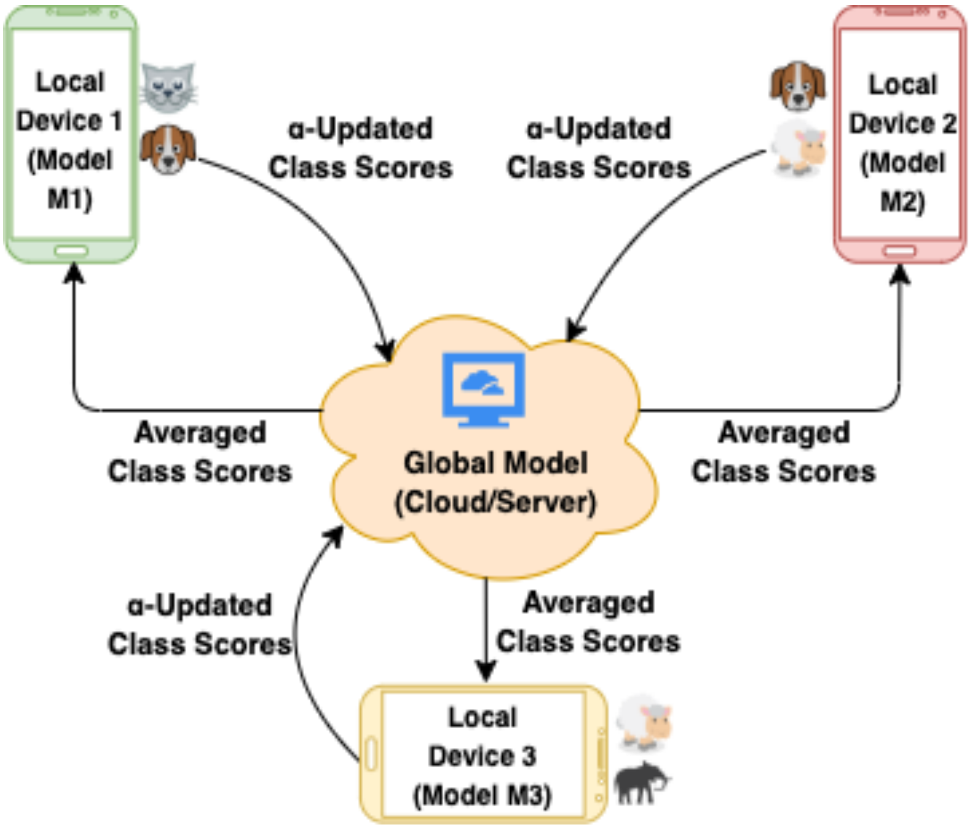
$$f_G^{i+1} = \sum_{m=1}^{M} \beta_m f_{\mathcal{D}_m^i}(x_0), \text{ where,}$$

$$\beta = \begin{cases} 1 & \text{If labels are unique} \\ \mathrm{acc}(f_{\mathcal{D}_m^i}(x_0)) & \text{If labels overlap} \end{cases}$$

  end for

end for

PROPOSED SYSTEM/ ARCHITECTURE



EXPERIMENTAL SETUP

- Animals-10 Dataset.
- 4 labels {Cat, Dog, Sheep, Elephant} simulated for 15 iterations across 3 users.
- D_0 is the public dataset (also test dataset), with 500 images per label – 2000 labels in total.
- Average the model scores on public dataset D_0 from the built models in each iteration.
- Image data across different iterations are split with disparities in both labels and distributions of data (*non-IID*).

	User 1	User 2	User 3	Global User
Model Arch.	2-Layer CNN {16, 32} Softmax Activation	3-Layer CNN {16, 16, 32} ReLU Activation	2-Layer CNN {16, 32} ReLU Activation	—
Labels	{Cat, Dog}	{Dog, Sheep}	{Sheep, Elephant}	{Cat, Dog, Sheep, Elephant}
Images per Iter	{500, 500}	{500, 500}	{500, 500}	{500, 500, 500, 500}

Alessio, C. Animals-10 Dataset, 2019. https://www.kaggle.com/alessiocorrado99/animals10

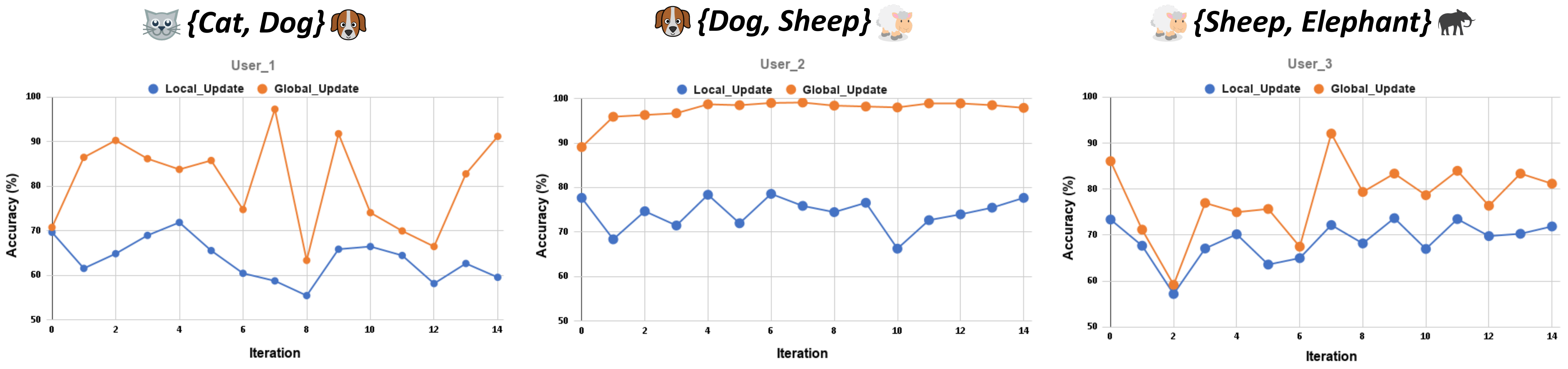
HETEROGENEITY IN MODEL ARCHITECTURES ACROSS ITERATIONS

Iterations	New Model Arch.
User 1 Iteration 10	3-Layer ANN (16, 16, 32) ReLU Activation
User 1 Iteration 14	1-Layer CNN (16) Softmax Activation
User 2 Iteration 6	3-Layer CNN (16, 16, 32) Softmax Activation
User 3 Iteration 5	4-Layer CNN (8, 16, 16, 32) Softmax Activation

AVERAGE INCREASE IN ACCURACIES ACROSS USERS

- Accuracies of all global updates in each user are deterministically higher than their respective accuracies of local updates.
- Information gain in User 2, maximum overlapped labels; more robust in global updates.
- Overall increase of **~16.7%** across all three users.

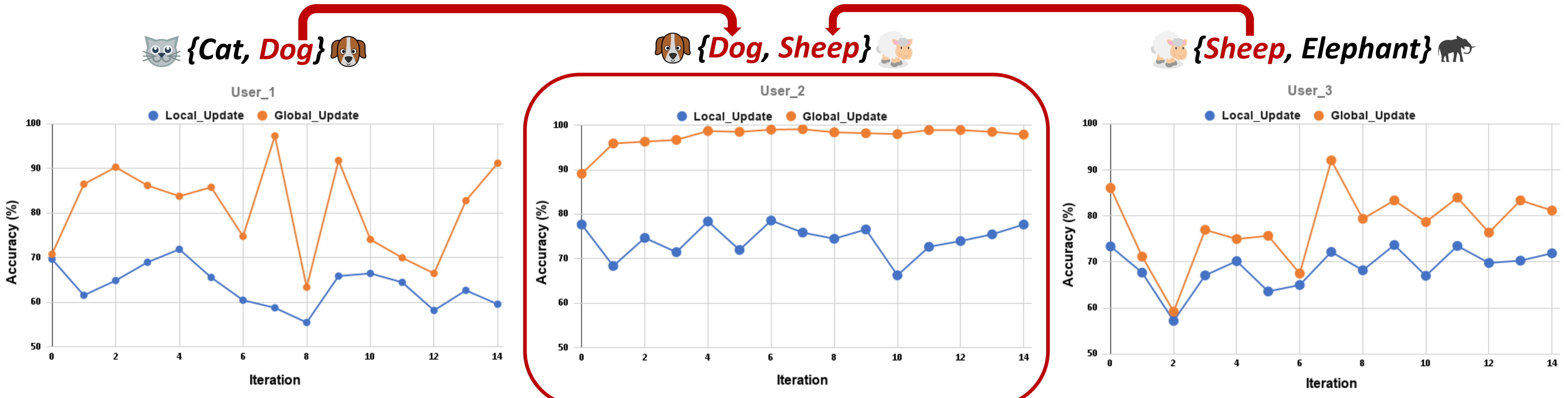
	Local Update	Global Update	Accuracy Increase
User 1	63.66	81.02	17.36
User 2	74.3	97.47	**23.17**
User 3	68.72	78.02	9.3
Average	**68.89**	**85.5**	**16.7**



LOCAL MODEL ACCURACY vs ITERATIONS

Local Update signifies the accuracy of each local updated model (after i^{th} iteration) tested on Public Dataset D_0.

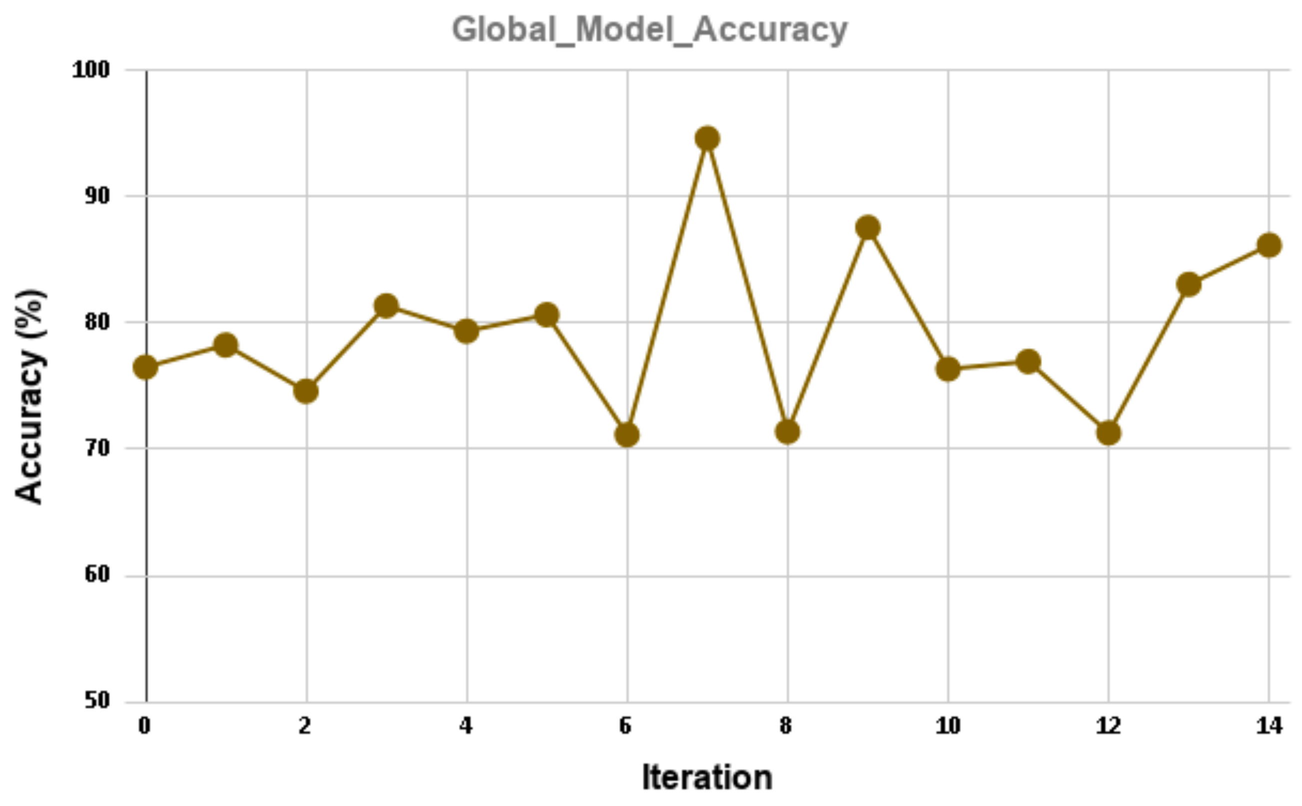
Global Update signifies the accuracy of the corresponding global updated model (after i^{th} iteration) tested on Public Dataset D_0.



LOCAL MODEL ACCURACY vs ITERATIONS

Local Update signifies the accuracy of each local updated model (after i^{th} iteration) tested on Public Dataset D_0.

Global Update signifies the accuracy of the corresponding global updated model (after i^{th} iteration) tested on Public Dataset D_0.



FINAL GLOBAL AVERAGE ACCURACIES vs ITERATIONS

ON-DEVICE PERFORMANCE

- On-device performance of our proposed framework is experimented on a Raspberry Pi 2.
- Similar (HW/SW) specifications with that of predominant contemporary IoT/edge/mobile devices.
- Clearly feasible.

Process	Computational Time
Training time per epoch in a FL iteration	1.8 sec
Inference time	15 ms

CONCLUSION

- A unified method with to handle both heterogeneous labels and model architectures in Federated Learning setting.
- Both global and local updates require computation of global model accuracy and are weighted based on it (α and β updates).
- Overlapping labels are found to make our framework robust, and also helps in effective accuracy increase.
- Exhibit on-device feasibility of federated learning and inference.

REFERENCES

[1] Li, T., Sahu, A. K., Talwalkar, A., and Smith, V. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine 37, 3 (2020), 50–60.

[2] McMahan, H. B., Moore, E., Ramage, D., Hampson, S., and Arcas, B. A., Communication-efficient learning of deep networks from decentralized data, In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (2017), vol. 54, pp. 1273–1282.

[3] Konečn`y, J., McMahan, H. B., Ramage, D., and Richtárik, P. Federated optimization: Distributed machine learning for on-device intelligence. arXiv preprint arXiv:1610.02527 (2016).

Contact

Gautham Krishna Gudur

Bala Shyamala Balaji

Perepu Satheesh Kumar

Let’s chat!

THANK YOU!

QUESTIONS?