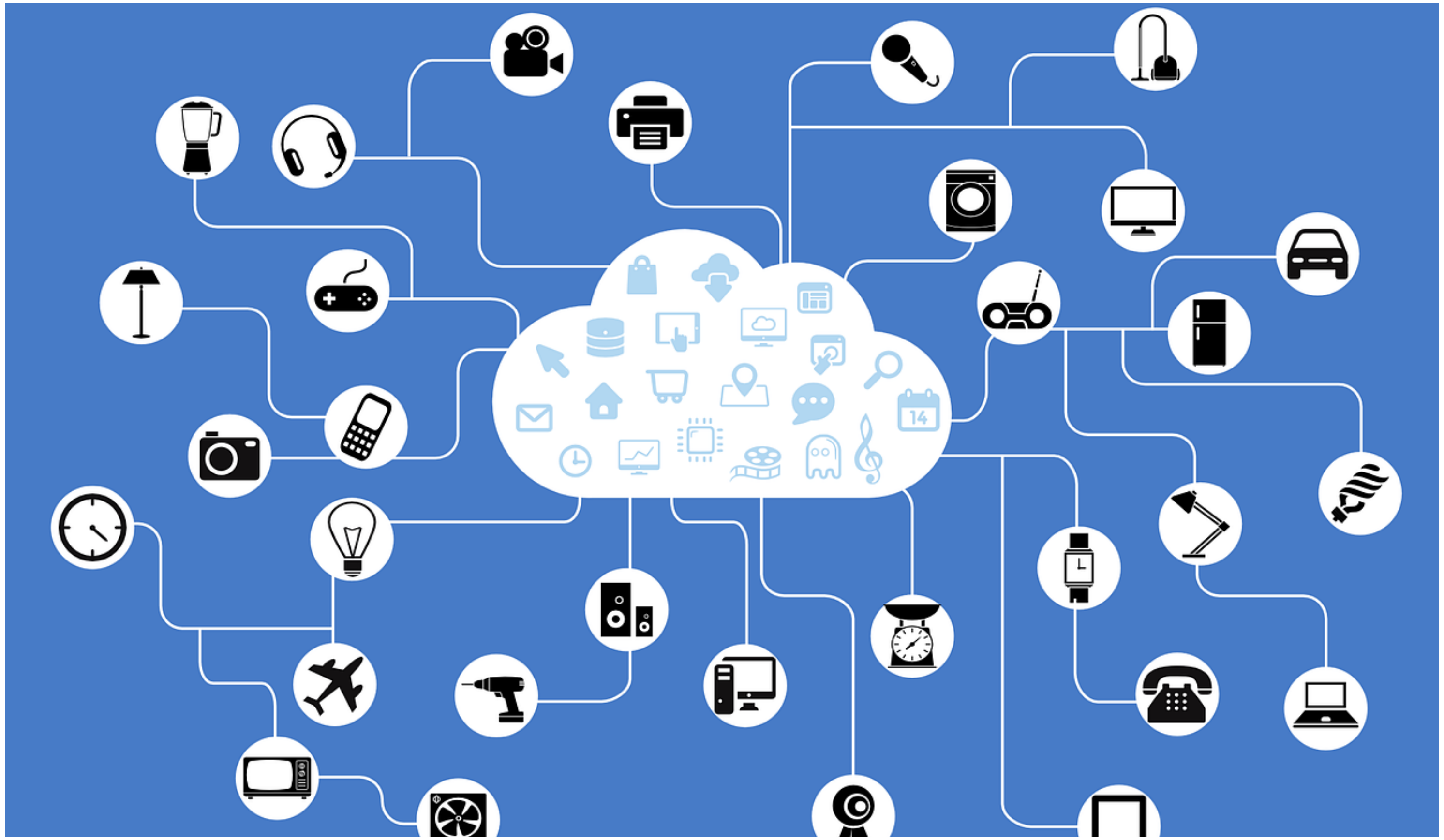
# RESOURCE-CONSTRAINED FEDERATED LEARNING WITH HETEROGENEOUS LABELS AND MODELS FOR HUMAN ACTIVITY RECOGNITION

**Gautham Krishna Gudur**
*Ericsson R&D*

**Perepu Satheesh Kumar**
*Ericsson Research*

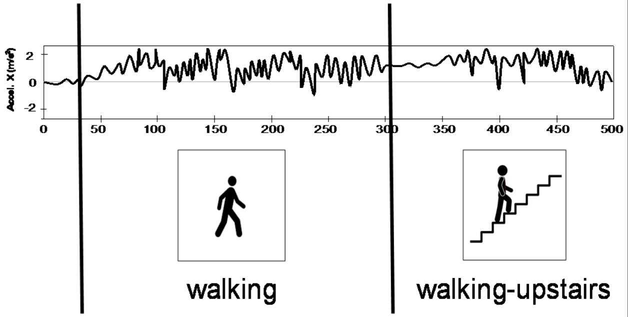


# IoT ON THE EDGE

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

An application particularly gaining importance in fields such as health-care and fitness tracking is

**Human Activity Recognition (HAR)**




# 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 |

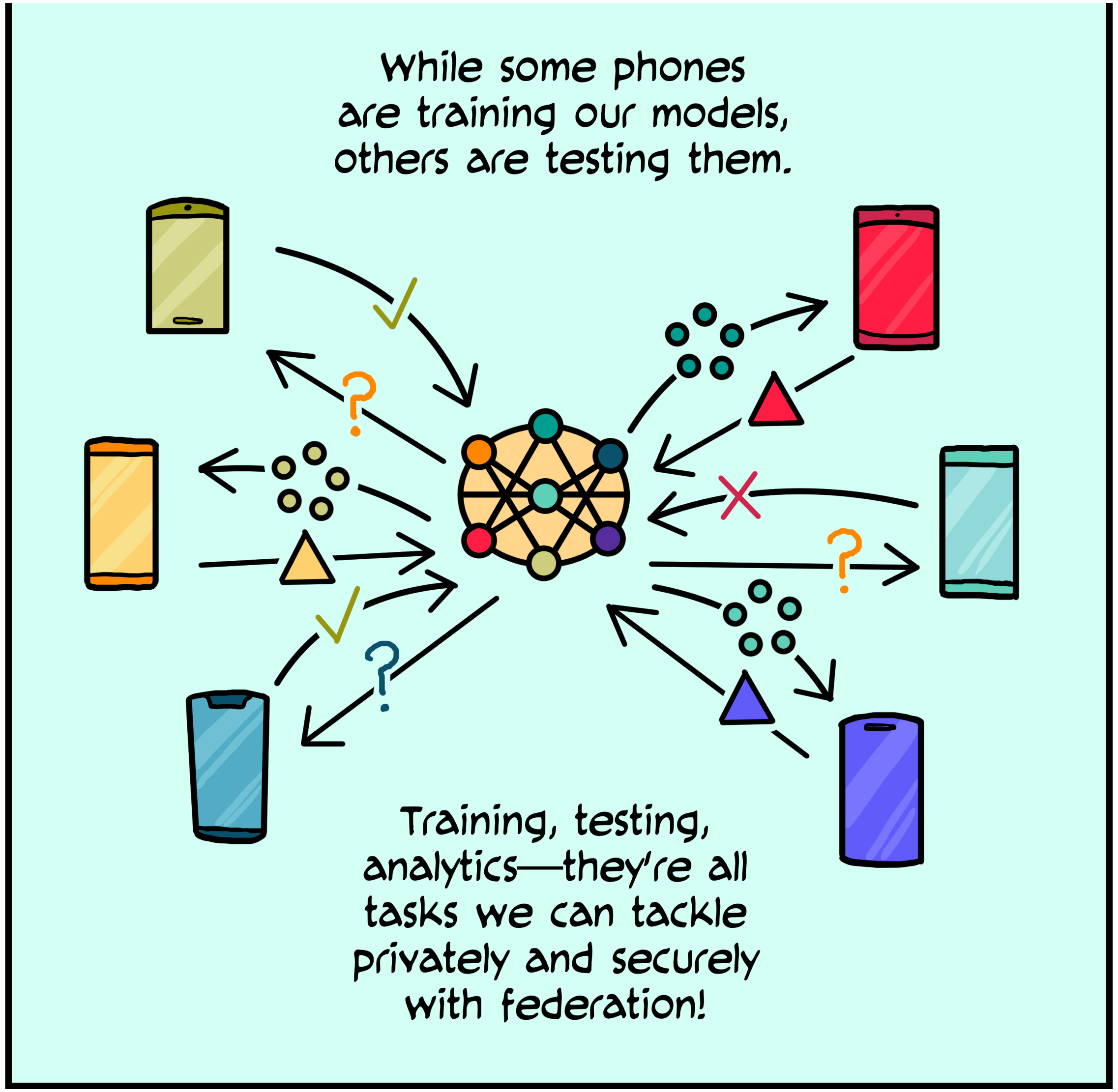
# FEDERATED LEARNING

Algorithms are trained across a federation of multiple decentralized devices.

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

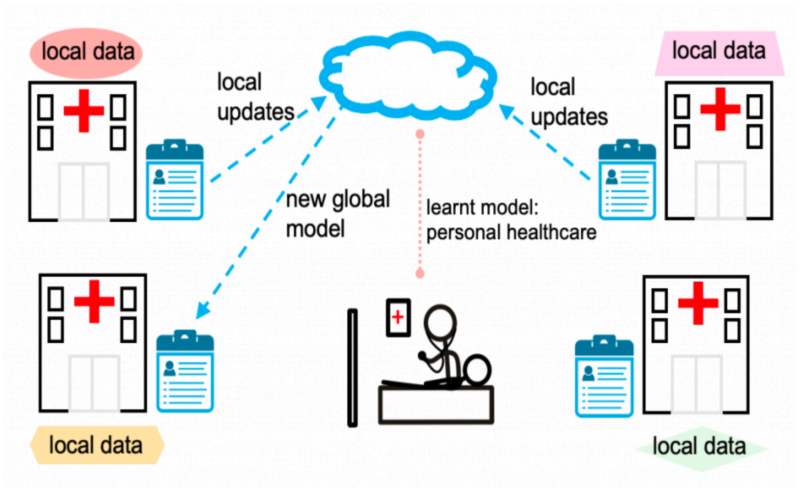
Transfer of model weights and updates from local devices to cloud, rather than conventional sharing of data.

***More Personalization; Minimal Latency; Privacy Preserving***.



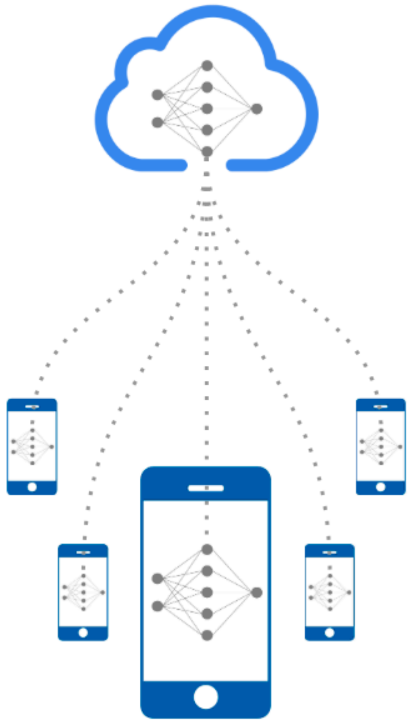

Picture taken from federated.withgoogle.com

## HEALTHCARE

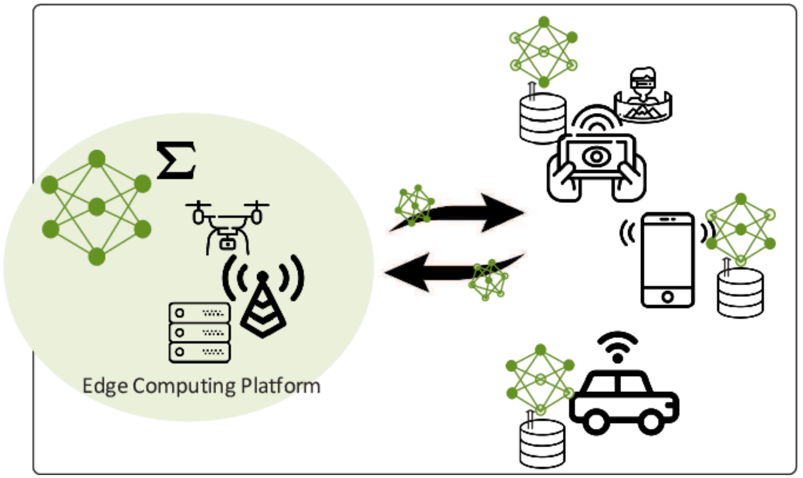



Picture taken from https://blog.ml.cmu.edu/2019/11/12/federated-learning-challenges-methods-and-future-directions/

## 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 (activities) across user devices.**

# GOALS OF OUR PROPOSED SYSTEM

- A framework to allow ***flexible heterogeneous selection of labels (activities)***, 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:*** To obtain scores across different iterations on a single user.
  - **Model Distillation Update:** Acts as summarization of models' information from older FL iterations.
  - **Weighted $\alpha$-update**: $\alpha$ - ratio between size of current private dataset and size of public dataset. Governs contributions of the new and old models.
- ***Global Update:*** Weighted average of scores across all users in same iteration.
  - $\beta$ parameter governs the weightage given to overlapping labels across users.

**Algorithm 1** Our Proposed Framework (with two version choices)

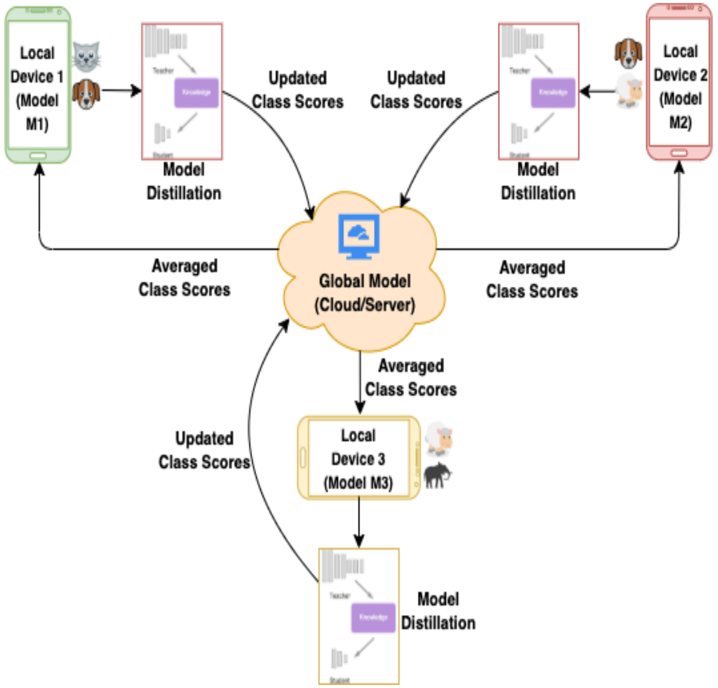
**Input:** Public Dataset $\mathcal{D}_0\{x_0, y_0\}$, Private Datasets $\mathcal{D}_m^i$, Total users $M$, Total iterations $I$, LabelSet $l_m$ for each user
**Output:** Trained Model scores $f_G^I$
Initialize $f_G^0 = \mathbf{0}$ (Global Model Scores)
**for** $i = 1$ **to** $I$ **do**
  **for** $m = 1$ **to** $M$ **do**
    **Build:** Model $\mathcal{D}_m^i$ and predict $f_{\mathcal{D}_m^i}(x_0)$
    **Local Update:**
    ***Choice 1 – Model Distillation Update:***
    Build a distilled model on only labels corresponding to local user's model with global averaged probabilities on public dataset $D_0$. Now, update the model with the new data $\mathcal{D}_m^i$ arriving in this iteration.
    ***Choice 2 – Weighted $\alpha$-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, $\alpha = \frac{len(\mathcal{D}_m^i)}{len(\mathcal{D}_0)}$
  **end for**
  **Global Update:** Update label wise

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

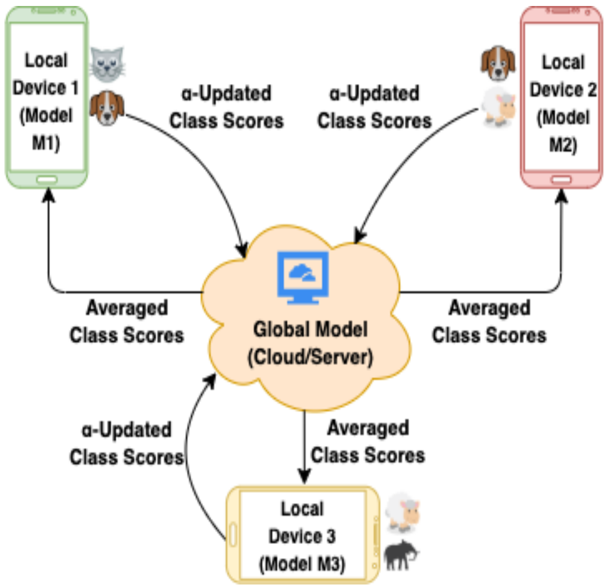
$$\beta = \begin{cases} 1 & \text{If labels are unique} \\ \text{acc}(f_{\mathcal{D}_m^{i+1}}(x_0)) & \text{if labels are not unique} \end{cases}$$

  where $\text{acc}(f_{\mathcal{D}_m^{i+1}}(x_0))$ is the accuracy function of the given model, and is defined by the ratio of correctly classified samples to the total samples for the given local model
**end for**

# PROPOSED SYSTEM/ ARCHITECTURE




**Model Distillation Update Version**




**Weighted-$\alpha$ Update Version**

# EXPERIMENTAL SETUP

- Heterogeneity Human Activity Recognition (**HHAR**) Dataset.
- Preprocessing: Discrete Wavelet Transform (DWT) and Decimation on accelerometer data; substantial data size decrease.
- 4 activities {Sit, Walk, Stand, StairsUp} simulated for 15 iterations across 3 users.
- $D_0$ is the public dataset (also test dataset), with 2000 activity windows per label – 8000 labels in total.
- Average the model scores on public dataset $D_0$ from the built models in each iteration.
- Activity 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 | 3-Layer ANN {16, 16, 32} ReLU Activation | — |
| **Activity Labels** | {Sit, Walk} | {Walk, Stand} | {Stand, StairsUp} | {Sit, Walk, Stand, StairsUp} |
| **Activity Windows per Iter** | {2000, 2000} | {2000, 2000} | {2000, 2000} | {2000, 2000, 2000, 2000} |

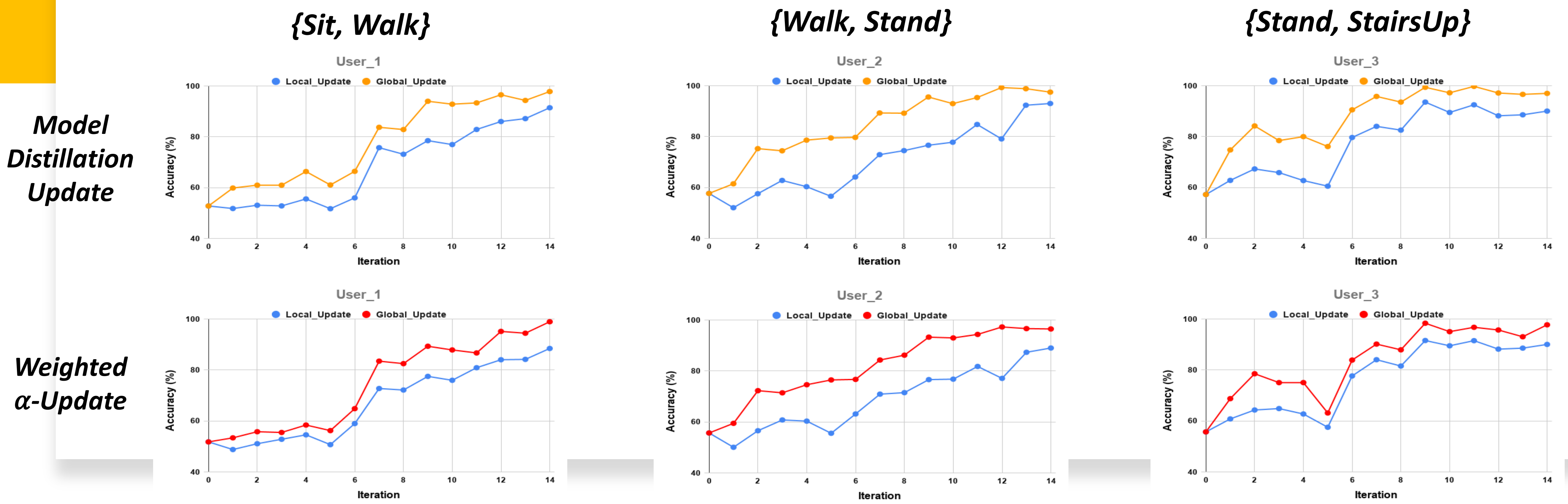
# HETEROGENEITY IN MODEL ARCHITECTURES ACROSS ITERATIONS

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

# AVERAGE INCREASE IN ACCURACIES ACROSS USERS

| | Model Distillation | | | Weighted $\alpha$-update | | |
|---|---|---|---|---|---|---|
| | **Local_Update** | **Global_Update** | **Increase** | **Local_Update** | **Global_Update** | **Increase** |
| **User_1** | 68.38 | 77.61 | 9.23 | 66.98 | 74.29 | 7.31 |
| **User_2** | 70.82 | 84.4 | 13.58 | 68.88 | 81.9 | 13.02 |
| **User_3** | 77.68 | 87.9 | 10.22 | 76.57 | 83.7 | 7.13 |
| **Average** | **72.293** | **83.303** | **11.01** | **70.81** | **79.963** | **9.153** |

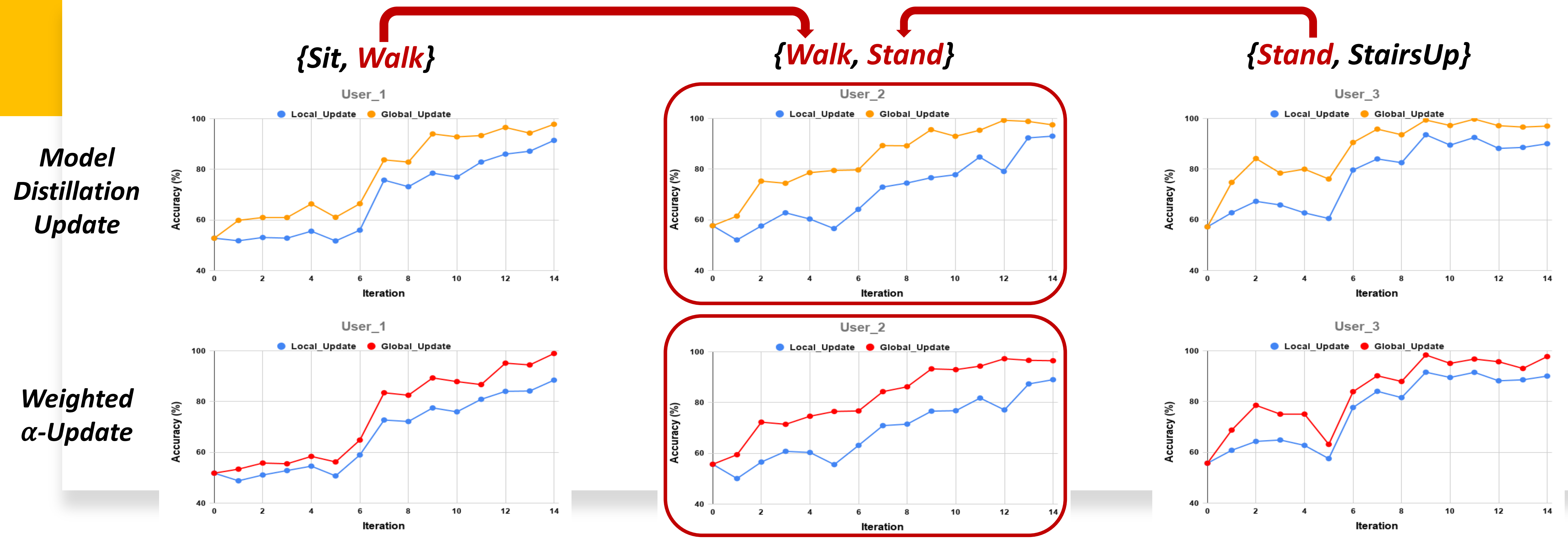
- Accuracies of all global updates in each user are deterministically higher than their respective accuracies of local updates for both proposed versions.
- Information gain in User 2, maximum overlapped labels; more robust in global updates.
- Overall increase,
  - **Model Distillation Update: ~11.01%** across all three users.
  - **Weighted $\alpha$-Update: ~9.153%** across all three users.




# LOCAL MODEL ACCURACY VS ITERATIONS

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

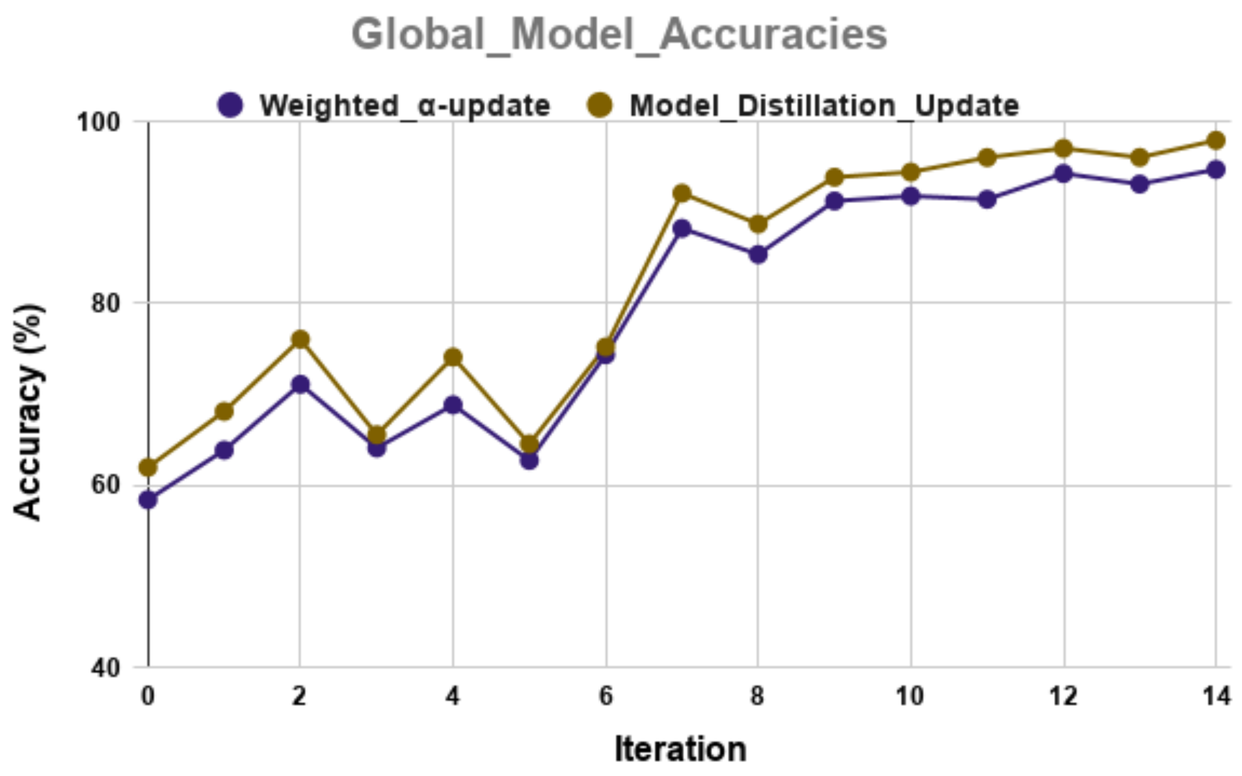
*Global Update* signifies the accuracy of the corresponding global updated model (after $i^{th}$ iteration) tested on Public Dataset $D_0$.




# 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 |
| Discrete Wavelet Transform (DWT) | 0.45 ms |
| Decimation | 4.6 ms |

# CONCLUSION

- A unified method with to handle both heterogeneous labels and model architectures in Federated Learning setting.
- Both global and local update accuracies are computed with two different local update versions – Model Distillation and Weighted $\alpha$-Update; and $\beta$ global update for label overlap.
- 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.

Contact

Gautham Krishna Gudur

Perepu Satheesh Kumar

Let’s chat!

# THANK YOU!

# QUESTIONS?