Flow: Per-instance Personalized Federated Learning

Kunjal Panchal[1] (kpanchal@umass.edu), Sunav Choudhary[2], Nisarg Parikh[3], Lijun Zhang[1], Hui Guan[1]
[1]University of Massachusetts – Amherst; [2]Adobe Research; [3]Northeastern University

NeurIPS 2023

TL;DR

- Flow is a per-instance and per-client personalization method to address the statistical heterogeneity issue in Federated Learning.
- Per-instance personalization addresses the following two shortcomings of Per-client personalization methods:
  - Performance of some clients is worse after personalization
  - For the clients who benefit from personalization, some instances still prefer the global model.
- Flow creates dynamic personalized models that are adaptive not only to each client's data distributions but also to each client's data instances.

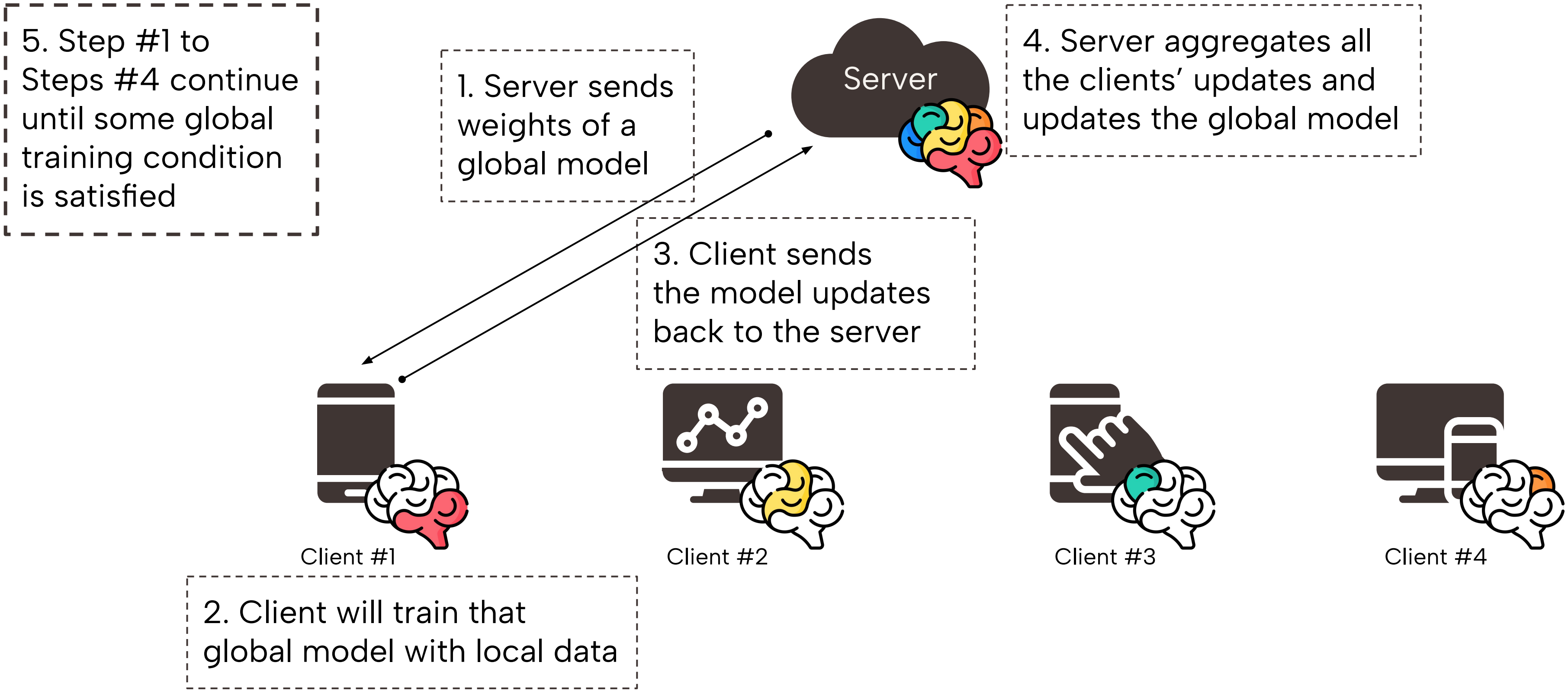
TL;DR

- Flow is a per-instance and per-client personalization method to address the statistical heterogeneity issue in Federated Learning.
- Per-instance personalization addresses the following two shortcomings of Per-client personalization methods:
    - Performance of some clients is worse after personalization
    - For the clients who benefit from personalization, some instances still prefer the global model.
- Flow creates dynamic personalized models that are adaptive not only to each client's data distributions but also to each client's data instances.

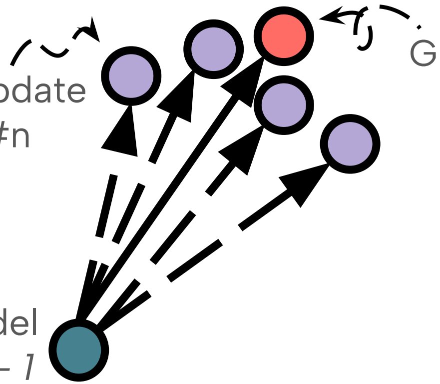
TL;DR

- Flow is a per-instance and per-client personalization method to address the statistical heterogeneity issue in Federated Learning.
- Per-instance personalization addresses the following two shortcomings of Per-client personalization methods:
  - Performance of some clients is worse after personalization
  - For the clients who benefit from personalization, some instances still prefer the global model.
- Flow creates dynamic personalized models that are adaptive not only to each client's data distributions but also to each client's data instances.

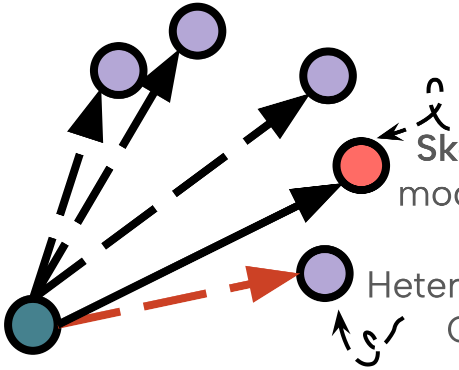
Federated Learning



The Data Heterogeneity Issue

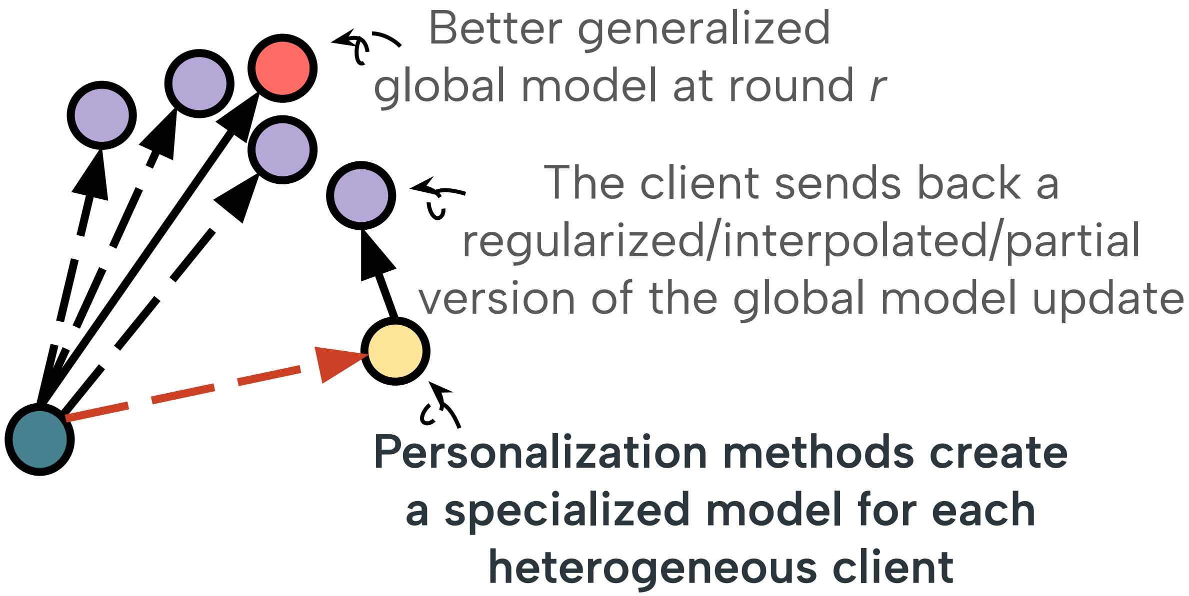


Clients with Homogeneous Data across Time



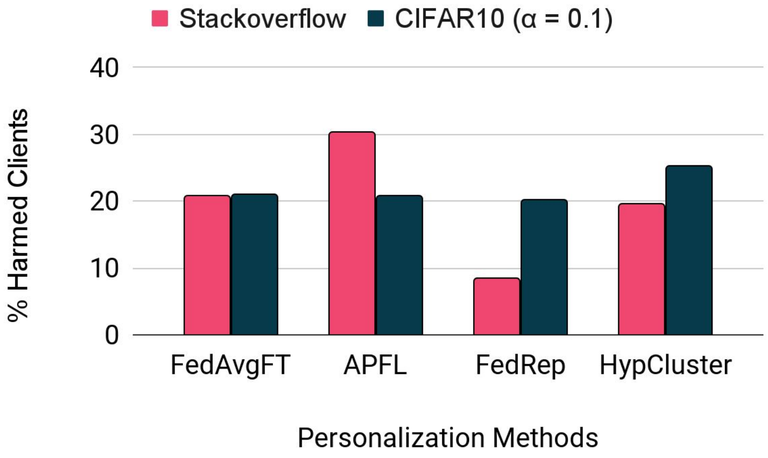
Different data distribution of a client with respect to other clients

Personalization to Rescue

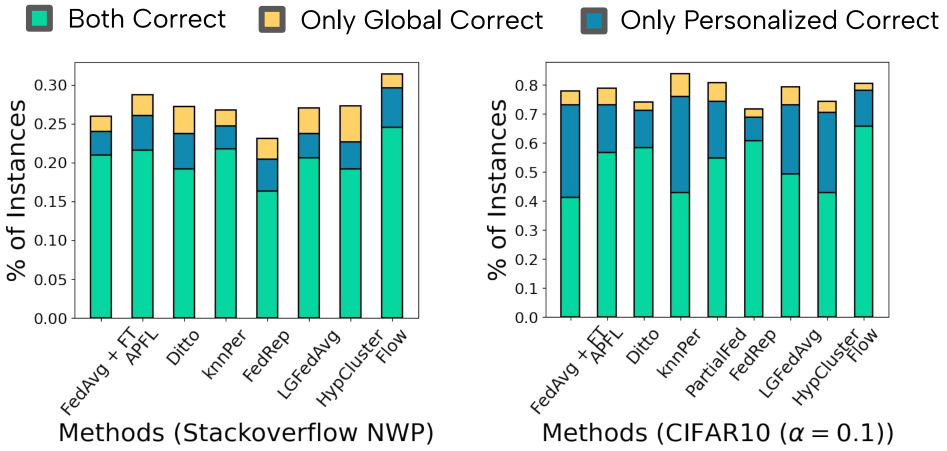


Client-wise Personalization is Limited

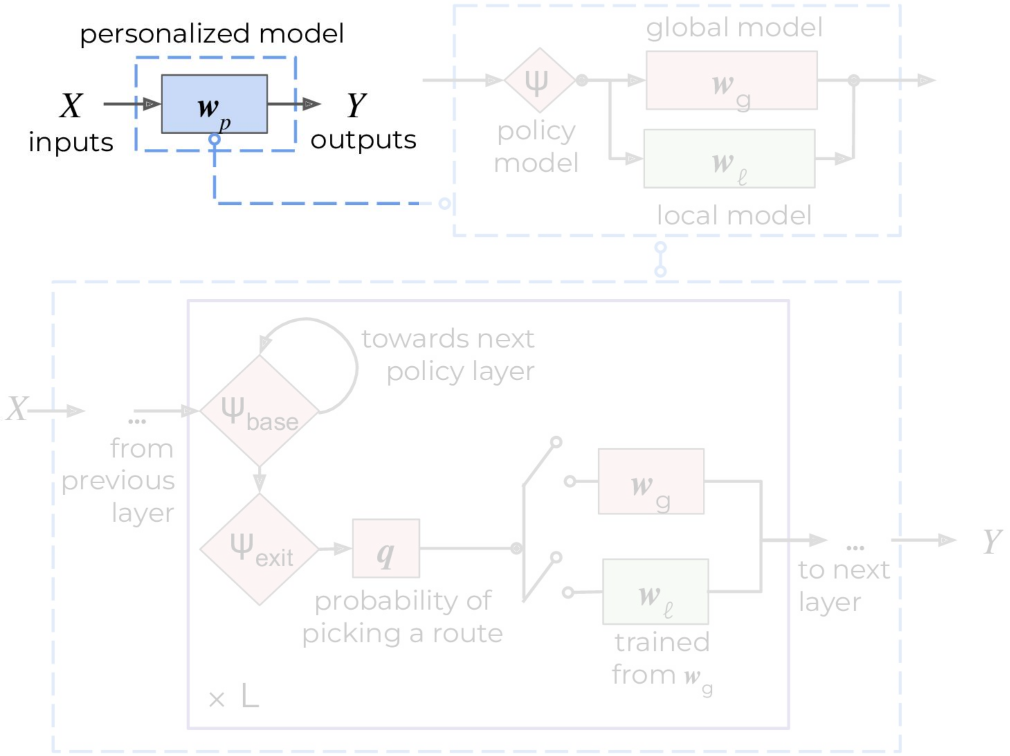
But, not all clients benefit from personalization



Even for the benefitting clients, not all instances prefer the personalized model

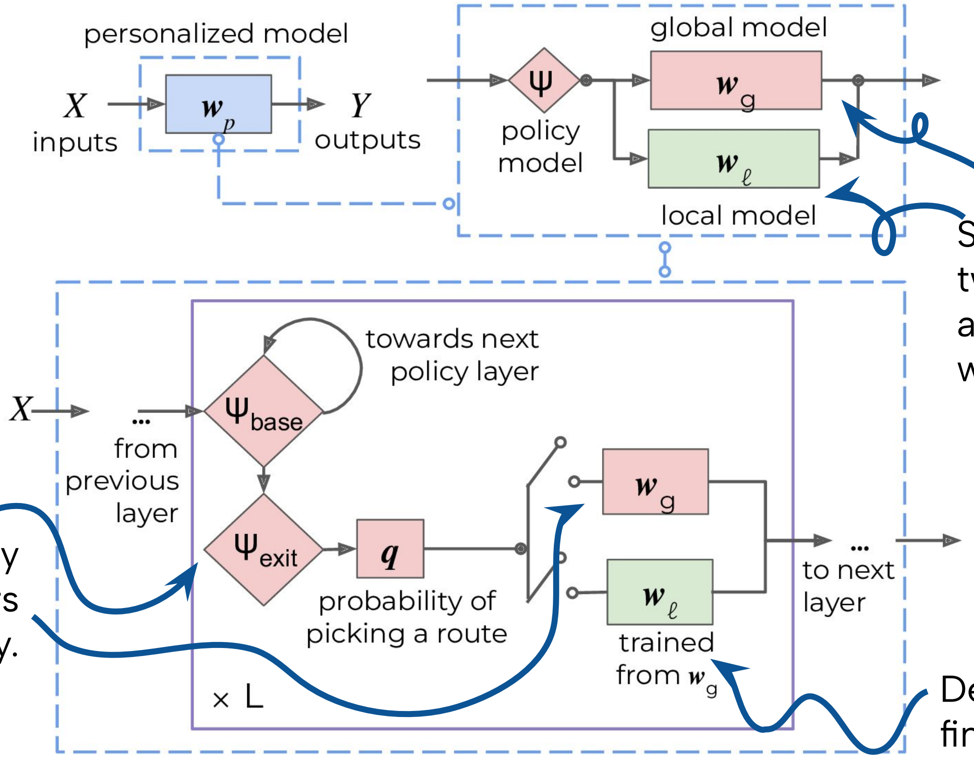


Flow



personalized model
X
inputs
w_p
Y
outputs
global model
Ψ
policy model
w_g
w_ℓ
local model
X
...
from previous layer
Ψ_{base}
towards next policy layer
Ψ_{exit}
q
probability of picking a route
w_g
w_ℓ
trained from w_g
...
to next layer
Y
× L

Flow

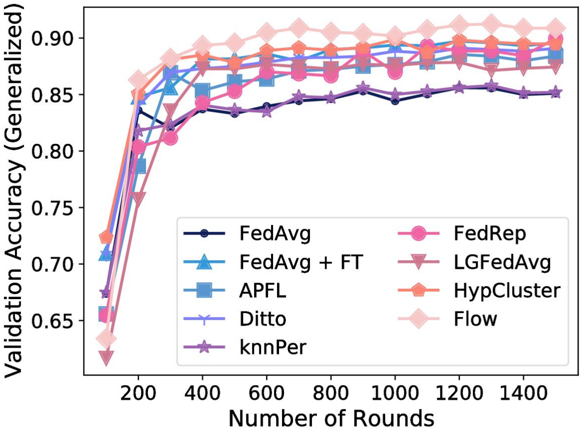


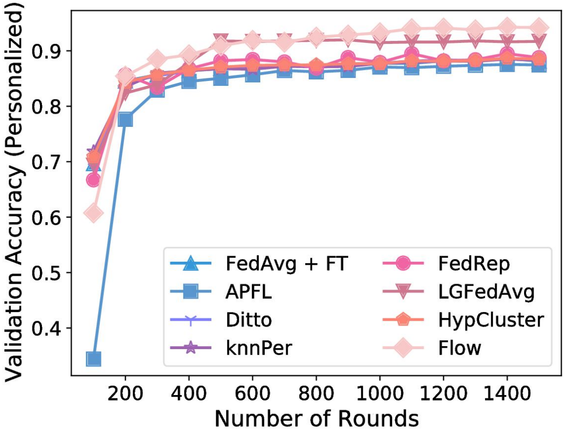
Split training dataset into two, for training global and local parameters with each half

Train the routing policy and global parameters alternatively.

Derive local parameters by finetuning the received global parameters

Results (Generalized and Personalized Accuracy)

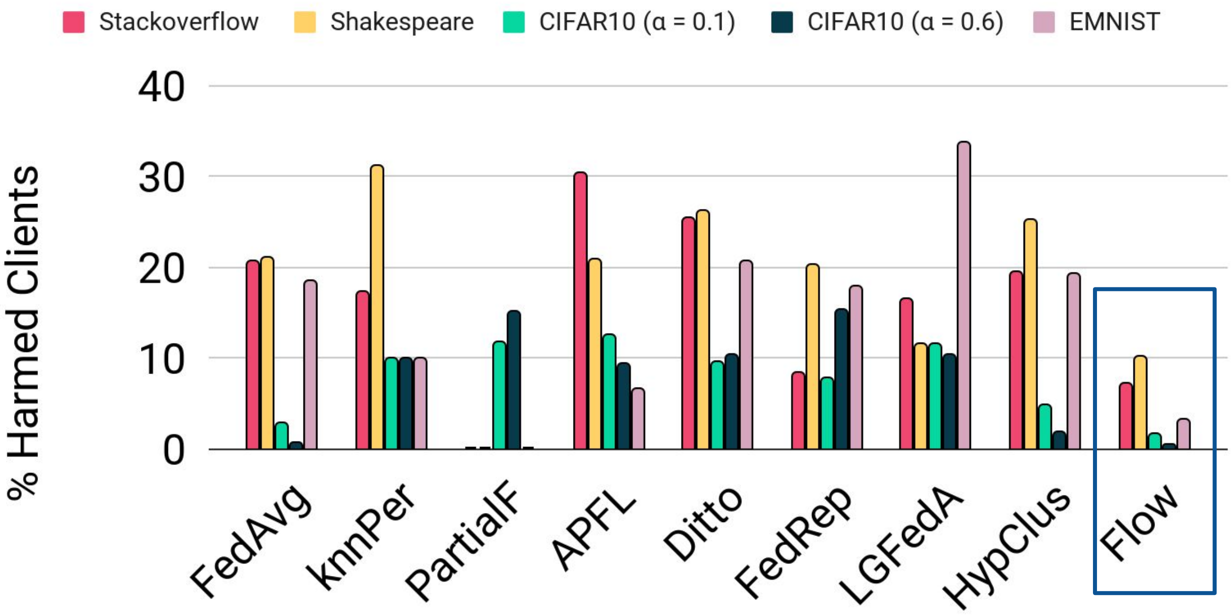




Observation 1
Flow achieves 1.11–3.46% higher generalized accuracy and 1.33–4.58% higher personalized accuracy over the best performing baseline.

Observation 2
Flow learns to put emphasis on data instances that are more aligned with the global data distribution to improve the performance of the global model.

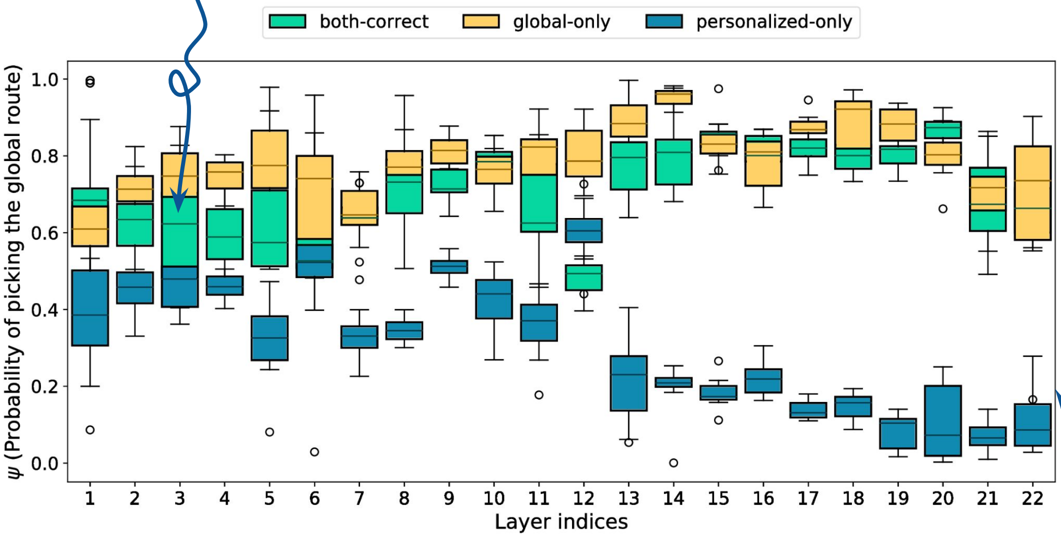
Results (Harmed Clients)



Observation 3
Flow achieves the highest percentage of clients who benefit from personalization compared to all personalization baselines, echoing the better personalized accuracy from Flow.

Results (Routing Policy Behavior)

For instances that can be correctly classified by both models (both-correct), the routing policy still prefers the global parameters over local parameters.



For instances that are correctly classified by w_g but not by w_p (global-only), we see a clear trend of the routing parameters getting more confident about picking the global parameters.

As a contrast, for instances that are correctly classified by w_p but not by w_g (personalized-only), we see the trend of routing policy being more confident in picking the local parameters.

Conclusion

- Flow creates dynamic personalized models with a routing policy that allow instances on each client to choose between global and local parameters to improve clients' accuracy.
- We derived error bounds for global and personalized models of Flow, showing how the routing policy affects the rate of convergence.
- The theoretical analysis validates our empirical observations related to clients preferring either a global or a local route based on the heterogeneity of individual instances.
- Extensive evaluation on both vision and language-based prediction tasks demonstrates the effectiveness of Flow in improving both the generalized and personalized accuracy.

Conclusion

- Flow creates dynamic personalized models with a routing policy that allow instances on each client to choose between global and local parameters to improve clients' accuracy.
- We derived error bounds for global and personalized models of Flow, showing how the routing policy affects the rate of convergence.
- The theoretical analysis validates our empirical observations related to clients preferring either a global or a local route based on the heterogeneity of individual instances.
- Extensive evaluation on both vision and language-based prediction tasks demonstrates the effectiveness of Flow in improving both the generalized and personalized accuracy.

Conclusion

- Flow creates dynamic personalized models with a routing policy that allow instances on each client to choose between global and local parameters to improve clients' accuracy.
- We derived error bounds for global and personalized models of Flow, showing how the routing policy affects the rate of convergence.
- The theoretical analysis validates our empirical observations related to clients preferring either a global or a local route based on the heterogeneity of individual instances.
- Extensive evaluation on both vision and language-based prediction tasks demonstrates the effectiveness of Flow in improving both the generalized and personalized accuracy.