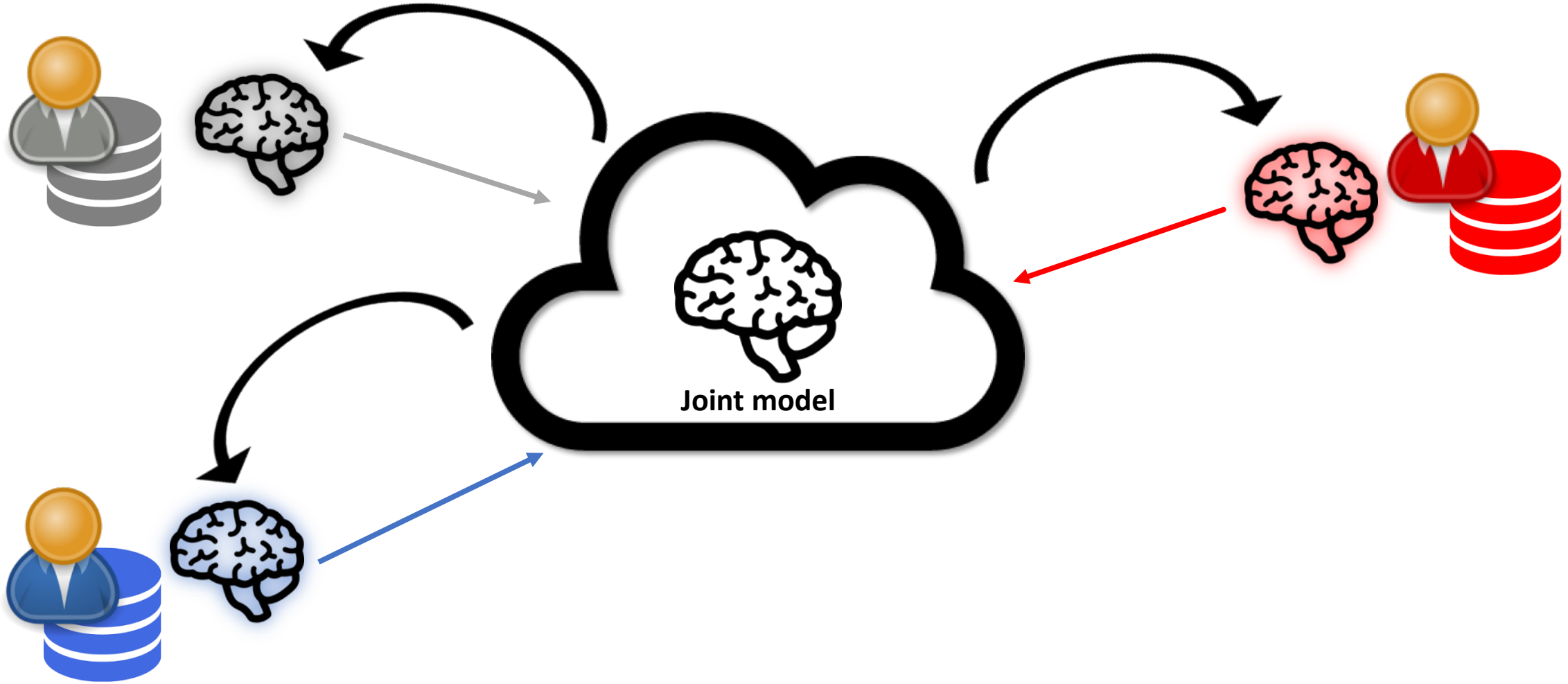
Collaborative/Federated Learning



Collaborative

Algorithm 1 Parameter server with synchronized SGD

Server executes:

Initialize θ_0
for $t = 1$ to T **do**
  for each client k **do**
    $g_t^k \leftarrow$ **ClientUpdate**(θ_{t-1})
  end for
  $\theta_t \leftarrow \theta_{t-1} - \eta \sum_k g_t^k$
end for

ClientUpdate(θ):

Select batch b from client's data
return local gradients $\nabla L(b; \theta)$

Federated

Algorithm 2 Federated learning with model averaging

Server executes:

Initialize θ_0
$m \leftarrow max(C \cdot K, 1)$
for $t = 1$ to T **do**
  $S_t \leftarrow$ (random set of m clients)
  for each client $k \in S_t$ **do**
    $\theta_t^k \leftarrow$ **ClientUpdate**(θ_{t-1})
  end for
  $\theta_t \leftarrow \sum_k \frac{n^k}{n} \theta_t^k$
end for

ClientUpdate(θ):

for each local iteration **do**
  for each batch b in client's split **do**
    $\theta \leftarrow \theta - \eta \nabla L(b; \theta)$
  end for
end for
return local model θ