l-Loss **B-Batch Size** n-Learning Rate  $C_1$ Client **Client**  $C_i$ - $i^{th}$  Cluster n- No of Clusters Client Client m-Size of i<sup>th</sup> Cluster  $C_2$  $W_m^o$ ,  $l_m$  $W^o$ ,  $n_m$ ,  $B_m$ Client  $W^o$ ,  $n_m$ ,  $B_m$  $W_m^o$ ,  $l_m$ Client Client Client  $W_m^o$ ,  $l_m$  $W^o, n_m, B_m$ Client Client  $W^o$ ,  $n_m$ ,  $B_m$  $C_0$  $W_m^o$ ,  $l_m$ (Broadcast) Genetic Algorithm(Cluster Wise) **Receive Hyperparameters** Client Client  $(n_m, B_m)$ Mutation Client Crossover Client Client Train Clients on **Evolve Hyperparameters**