



(51) International Patent Classification:
G06N 20/20 (2019.01)

(21) International Application Number:

PCT/US2023/032687

(22) International Filing Date:

14 September 2023 (14.09.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/407,530 16 September 2022 (16.09.2022) US
63/408,553 21 September 2022 (21.09.2022) US
18/466,333 13 September 2023 (13.09.2023) US

(71) Applicant: **NEC LABORATORIES AMERICA, INC.**
[US/US]; 4 Independence Way, Suite 200, Princeton, New
Jersey 08540 (US).

(72) Inventors: **CHENG, Wei**; 5 Arnold Drive, Princeton Junc-
tion, New Jersey 08550 (US). **YU, Wenchao**; 2110 Fox
Run Drive, Plainsboro, New Jersey 08536 (US). **CHEN,**

Haifeng; 11 Blackhawk Court, West Windsor, New Jersey
08550 (US). **WU, Yue**; 660 Veteran Avenue, Los Angeles,
California 90024 (US).

(74) Agent: **BITETTO, James J.**; 401 Broadhollow road, Suite
402, Melville, New York 11747 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,
KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA,
NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,
RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, CV,
GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST,

(54) Title: PERSONALIZED FEDERATED LEARNING UNDER A MIXTURE OF JOINT DISTRIBUTIONS

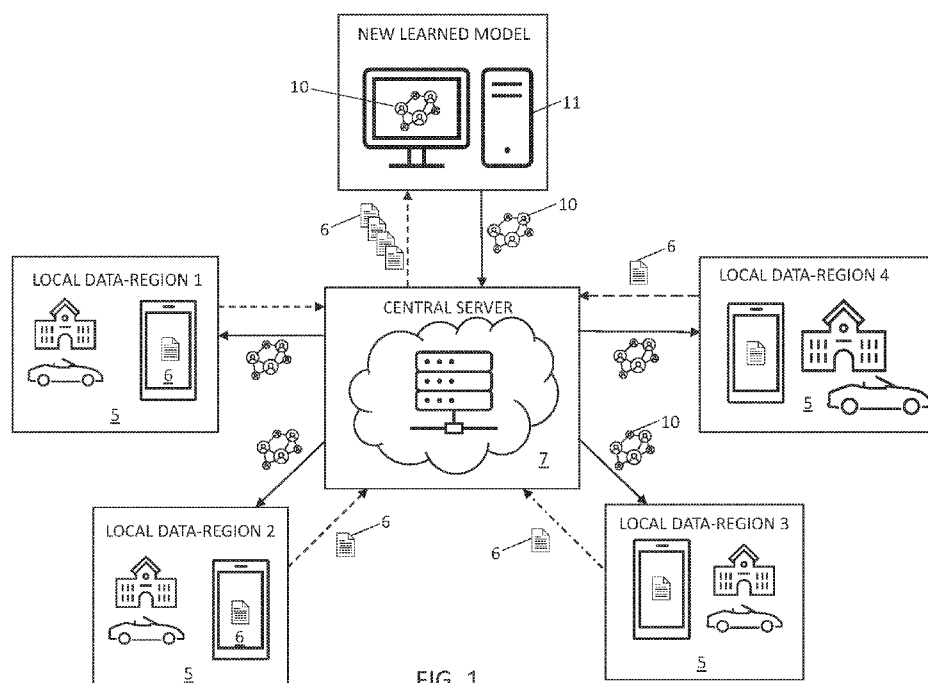


FIG. 1

(57) Abstract: Systems and methods for personalized federated learning. The method may include receiving at a central server (7) local models (6) from a plurality of clients (5), and aggregating a heterogeneous data distribution extracted from the local models (6). The method can further include processing the data distribution as a linear mixture of joint distributions to provide a global learning model (10), and transmitting the global learning model (10) to the clients (5). The global learning model (10) is used to update the local model (5).

SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

PERSONALIZED FEDERATED LEARNING UNDER A MIXTURE OF JOINT DISTRIBUTIONS

RELATED APPLICATION INFORMATION

[0001] This application claims priority to US 63/407,530 filed on September 16, 2022, incorporated herein by reference in its entirety.

[0002] This application claims priority to US 63/408,553 filed on September 21, 2022, incorporated herein by reference in its entirety.

[0003] This application claims priority to U.S. Patent Application Serial No. 18/466,333 filed on September 13, 2023, incorporated herein by reference in its entirety.

BACKGROUND

Technical Field

[0004] The present invention relates to machine learning and more particularly to federated learning.

Description of the Related Art

[0005] When moving to the high level human needs, the information that is used in machine learning is getting more personal. For example, banking information is personal. The detailed home address and social security number (SSN) of an individual

are also personal information. Private chats in social networks can be personal. The security of this information can be critical to the safety and esteem of individuals.

[0006] As more attention is being put on the high-level needs using artificial intelligence and machine learning, among the most prominent concerns people have about modern technologies that access consumer/client data are those regarding privacy. It is common to assume that privacy is all about being able to control information about oneself, generally in the form of having a right to prevent others from obtaining or using information about a person without that person's consent. People have more and more interests in controlling information about themselves or preventing others from knowing things about themselves without their consent. Protection of such "security interests" constitutes an important reason people have for wanting to avoid having others acquire information about them.

[0007] Another factor to consider is the personalization. As we all know, premium services, such as those employing artificial intelligence and machine learning, are usually personalized and customized. In today's market, customer service is a huge differentiator. In order to earn loyalty, business needs to deliver personalized customer service. This means more than merely satisfying their low-level basic needs. Delivering truly personalized customer service is quite a feat. It entails making the customer feel like they are dealing with a company which treats them humanely. Protection of private data is the utmost concern for some consumers utilizing high level services, such as machine learning.

SUMMARY

[0008] According to an aspect of the present invention, a computer implemented method is provided for personalized federated learning. In one embodiment, the method can include receiving at a central server local models from a plurality of clients; and aggregating a heterogeneous data distribution extracted from the local models. In one embodiment, the method can further include processing the data distribution as a linear mixture of joint distributions to provide a global learning model; and transmitting the global learning model to the clients, wherein the global learning model is used to update the local model.

[0009] In accordance with another embodiment of the present disclosure, a system for personalized federated learning is described that includes a hardware processor; and memory that stores a computer program product. The computer program product when executed by the hardware processor, causes the hardware processor to receive, using the hardware processor, at a central server local models from a plurality of clients; and aggregate, using the hardware processor, a heterogeneous data distribution extracted from the local models. The computer program product can also process, using the hardware processor, the data distribution as a linear mixture of joint distributions to provide a global learning model; and transmit, using the hardware processor, the global learning model to the clients, wherein the global learning model is used to update the local model.

[0010] In accordance with yet another embodiment of the present disclosure a computer program product for personalized federated learning. The computer program product can include a computer readable storage medium having computer readable

program code embodied therewith. The program instructions executable by a processor to cause the processor to receive, using the hardware processor, at a central server local models from a plurality of clients; and aggregate, using the hardware processor, a heterogeneous data distribution extracted from the local models. In some embodiments, the program instructions can also include to process, using the hardware processor, the data distribution as a linear mixture of joint distributions to provide a global learning model; and transmit, using the hardware processor, the global learning model to the clients, wherein the global learning model is used to update the local model.

[0011] These and other features and advantages will become apparent from the following detailed description of illustrative embodiments thereof, which is to be read in connection with the accompanying drawings.

BRIEF DESCRIPTION OF DRAWINGS

[0012] The disclosure will provide details in the following description of preferred embodiments with reference to the following figures wherein:

[0013] FIG. 1 is a illustration of a general environment depicting federated learning, in accordance with one embodiment of the present disclosure.

[0014] FIG. 2 is a block/flow diagram of an exemplary method that uses a Federated Gaussian Mixture Model (Fed-GMM) to model the joint probability of samples in each client.

[0015] FIG. 3 is a block/flow diagram of an exemplary method for local training of the edge devices using local data, in accordance with embodiments of the present invention.

[0016] FIG. 4 is a block/flow diagram of an exemplary method for sending parameters to a central server in the federated learning environment, in accordance with embodiments of the present invention.

[0017] FIG. 5 is a block/flow diagram of an exemplary method aggravating parameters from a Gaussian mixture model, in accordance with embodiments of the present invention.

[0018] FIG. 6 is a block/flow diagram of an exemplary processing system for personalized federated learning, in accordance with embodiments of the present invention.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

[0019] In accordance with embodiments of the present invention, systems and methods are provided for personalized federated learning. Personalized customer experiences employing machine learning and artificial intelligence may leverage customer data, thus privacy concerns arise.

[0020] Enterprises have been using model-centric machine learning (ML) approaches. In this example, a central machine learning (ML) model can be employed, which is present in the central server, and is trained with all available training data. With the recent data-centric artificial intelligence (AI) trend, the focus of artificial intelligence (AI) is shifting from model-centric to data-centric approaches. Data-centric artificial intelligence (AI) focuses on systematically improving the available data, which yields faster model building, shorter time for deployment, and improved accuracy. However,

this approach brings its own set of challenges regarding the accessibility of the data. One problem is getting access to those huge amounts of data without causing data breaches to the companies or to individuals.

[0021] In response to rising concerns about data privacy, federated learning has been instituted, which allows training models collaboratively without sharing raw local data. This method brings the model to the local data rather than gathering the data in one place for the model training.

[0022] Referring to FIG. 1, the principle of federated learning is very simple. All clients 5 that have data on their devices, such as smartphones. A smartphone is only one example of a data containing device that may be used by the clients 5, as the data for the clients 5 may also be generated by sensor data from cars, branches of bank, and hospitals. The data is used at the individual clients 5 to train their individual models 6 (also referred to as local models 6). The individual model is identified by reference number 6. The clients 5 then send the model 6, not the data, to a central server 7 that combines the models for each of the individual classes and sends out the new combined model 8 to every client for further updating rounds. The central server 7 can aggregate the data. The aggregated data may then be used by a model generator 9 to create a new global model 10.

[0023] As noted, privacy interests have to be considered. To address the issue, the model generator 9 performs personalized federated learning (PFL), which collaboratively trains a federated model, i.e., generates a new global model 10, while considering local clients 5 under privacy constraints. It has been observed that existing PFL approaches

result in suboptimal solutions when the joint distribution among local clients 5 diverges. The clients 5 can be highly different. For example, they can have different input distribution; they can have different conditional distribution; and the number of samples can be different.

[0024] In one example, a car sales prediction is considered, in which there are several dealers, e.g., sever auto sales dealers around the world, who want to use machine learning to predict how likely a car will be sold, as illustrated in FIG. 1. In some embodiments, the prediction model can take into account the car's specs and price as input and predict how likely the customer will buy it. In some embodiments, the model, make, color and price distributions of cars in the market are different in each regions (e.g., local data region 1 vs. local data region 2), which means the input distribution will be different. Second, even if the car and its price are the same, depending on the culture and economy, people from different regions can still have different preferences. This means the conditional distribution is also different, e.g., an example of being heterogeneous. A personalized algorithm added to the model is expected to both utilize the data globally and adapt to different markets. These heterogeneities indeed post obstacles for existing methods. Either they make some unrealistic assumptions, or propose very specific methods that are not generalizable to new tasks. It is noted that in this example, the prediction model is an example of a local model 6, as depicted in FIG. 1. The aforementioned make, model, price, color and local regions are all examples of parameters of the local model. These parameters can be extracted from the local model without divulging private information of the clients 5.

[0025] The methods, systems and computer program products that are described herein provide personalized federated learning under a mixture of joint distributions. In machine learning, a “distribution” is simply a collection of data, or scores, on a variable. In some embodiments, these scores are arranged in order from smallest to largest. Distribution can refer to how the data is spread out or clustered around certain values or ranges. By examining the distribution, insights can be gained into the characteristics and patterns of the data, which can be useful in making informed decisions and predictions

[0026] In machine learning, joint distributions refer to the probability distribution of two or more variables occurring together. These variables could be any features or attributes that are related to a particular problem.

[0027] In some embodiments, the goal of the methods, systems and computer program products that are described herein for personalized federated learning is to model the joint probability as a whole, and address the heterogeneity issue in a formal, rigorous way. Heterogeneity means that there is variability in the data. The opposite of heterogeneity is homogeneity meaning that all studies show the same effect.

[0028] Federated learning (FL) is a framework that allows edge device to collaboratively train a global model while keeping customers' data on-device. A standard FL system involves a cloud server 7 and multiple clients (devices). Each device has its local data and a local copy of the machine learning (ML) model 6 being served.

[0029] In each round of federated learning (FL) training, a cloud server 7 sends the current global model 10 to the clients 5, and the clients train 5 their local models 6 using on-device data and send the models 6 to a centralized server 7 (which may be cloud based

4). The server 7 aggregates the local models 6 and updates the global model 10. In some embodiments, the Federated learning (FL) also has a personalization branch, which aims to customize local models to improve their performance on local data.

[0030] Traditional federated learning (FL) assumes that clients are homogenous. However, in reality the data distribution $\mathbb{P}(x, y)$ of clients can be heterogenous. A personalized federated learning (FL) algorithm aims to utilize data from all clients 6, and generate model, e.g., global models 10, suitable for different clients.

[0031] In some embodiments, a Federated Gaussian Mixture Model (Fed-GMM) is proposed to jointly model the joint probability of samples in each client 5. This approach can use the linear mixture of several base distributions, while the weights π_c is personalized for each client. Given the model of the likelihood, we use log-likelihood maximization as the training criterion. The proposed model can automatically detect if the new clients data are from the same joint distribution of the data in existing clients. Thus, it is able to detect outlier of clients.

[0032] In one example, the data distribution $\mathbb{P}(x)$ is data includes colors, makes, models and pricing in the market for vehicle sales. The global learned model may provide based upon a given price, etc., a consumer's willingness to buy. This is only one example of what can be modeled by the global learned model, which can result from the aggregated data from local models from the clients in a federated learning scenario. As noted prior to the methods of the present disclosure, assume that the data is homogenous, and fail to address client heterogeneity. In some examples, the objective of the present disclosure addresses joint distribution heterogeneity in federated learning models.

[0033] In one embodiment, the methods, systems and computer program products of the present disclosure can model $\mathbb{P}(x,y)$ as a linear mixture of joint distributions. In one embodiment, the formulation:

$$\text{Equation (1):} \quad \mathbb{P}_c(x,y) = \sum_{i=1}^M \pi_c^i \mathbb{P}^i(x,y), \mathbb{P}^i(x,y)$$

[0034] Can be applied to deep networks, in which the optimization goal, which is a log likelihood maximization, can be provided by:

$$\text{Equation (2):} \quad \forall c \in C, \quad \max_{\pi_c, \theta} \mathbb{E}_{(x,y) \sim \mathbb{P}_c} [\log(\mathbb{P}_{\pi_c, \theta}(x,y))]$$

[0035] In accordance with Equation (2), for a fixed client c , the methods, systems and computer program products of the present disclosure consider a fixed client c , which has a defined latent variable that is equal to:

$$\text{Equation (3):} \quad z \sim \pi_c(\cdot), \text{ then } (x,y) \sim \mathbb{P}(\cdot | z) = \mathbb{P}_{\theta^z}(\cdot)$$

[0036] in which the definition in Equation (3), implies:

$$\text{Equation (4):} \quad \mathbb{P}_{\pi_c, \theta}(x,y) = \sum_{i \in |M|} \pi_c^i \mathbb{P}_{\theta^i}(x,y)$$

[0037] The approach for the federated learning model ($\mathbb{P}(x,y)$) that employs a linear mixture of joint distributions with Equations (1), (2), (3) and (4), can advantageously provide a highly expressive model, in which an increase in M provides for higher expressive power. The federated learning model ($\mathbb{P}(x,y)$) that employs a linear mixture of joint distributions with Equations (1), (2), (3) and (4), can advantageously be used to better guide supervised learning. Even further, the federated learning model ($\mathbb{P}(x,y)$) that

employs a linear mixture of joint distributions with Equations (1), (2), (3) and (4), can advantageously be used for out of distribution detections. The equations described above with reference to equations (1)-(4), provides a simple yet powerful formulation for a personalized federated learning (FL) task.

[0038] The federated learning algorithm described above with reference to FIGs. 1-4 provides better expressivity and better supervised learning guidance. For example, existing federated learning models prior to the methods and systems of the present disclosure assumed a homogenous data distribution, i.e., assumed $\mathbb{P}(x)$ to be the same. One example of a prior learning model is as follows:

Equation (5):
$$\mathbb{P}(y|x) = \pi_1 \mathbb{P}_1(y|x) + \pi_2 \mathbb{P}_2(y|x)$$

[0039] Equation (5) assumes that $\mathbb{P}(x)$ to be the same. The resulting classifier suffers from constant errors due to the assumption that $\mathbb{P}(x)$ is the same. The model provided by the methods, systems of the computer program products of the present disclosure does not assume homogenous data distributions. Contrary to the modeling provided by Equation (5), a federated learning model as a mixture of joint distributions to address data distributions $\mathbb{P}(x,y)$ that are heterogeneous is provided by Equation (6):

Equation (6):
$$\mathbb{P}(y|x) = \frac{\pi_1 \mathbb{P}_1(x) \mathbb{P}_1(y|x) + \pi_2 \mathbb{P}_2(x) \mathbb{P}_2(y|x)}{\pi_1 \mathbb{P}_1(x) + \pi_2 \mathbb{P}_2(x)}$$

[0040] The federated learning model provided by equation (6) can express a Bayesian optimal classifier.

[0041] The model provided by equation 6 is a federated gaussian mixture model (GMM), which can be used to estimate the probability ($\mathbb{P}_c(x)$) that a client (c) will act based upon the data (x). The base model is a Gaussian distribution $\mathbb{P}^i(x, y) = \mathcal{N}(x; \mu_i, \Sigma_i) \mathbb{P}_{\theta^i}(y|x)$. The $\mathbb{P}_{\theta^i}(y|x)$ can be a deep supervised model parameterized by θ^i . Algorithm derivation for the federated gaussian mixture model may include the following.

[0042] The general model may employ $[N]$ to denote the set $\{1, 2, \dots, N\}$, e.g., set of data. Suppose there exist C clients, e.g., the illustration in FIG. 1 illustrates four (4) clients 10. In some embodiments, each client $c \in [C]$ has its own dataset of size N_c , where a data $s_{c,i} = (x_{c,i}, y_{c,i})$ is assumed to be drawn from distribution $\mathbb{P}_c(x, y)$.

[0043] The model $\mathbb{P}(x, y)$ of the computer implemented methods, systems computer program products of the present disclosure employs a linear mixture of joint distribution

[0044] The multi-task federated learning tasks require a model $P_c(x, y)$ (identified by reference number 6 in FIG. 1) for each client 10, such that they can achieve expected maximum log likelihood: $\forall c \in \mathcal{C}, \max_{\pi_c, \theta} \mathbb{E}_{(x, y) \sim \mathbb{P}_c} [\log (P_c(x, y))]$.

[0045] In some embodiments, linear mixtures can be employed to model the probability. For the input distribution $P_c(x)$, Gaussian mixture models (GMM) are employed. For the conditional distribution $P_c(y | x)$, parameterized supervised learning models are employed.

[0046] The model is defined as follows:

[0047] All clients share the GMM parameters $\{\mu_{m_1}, \Sigma_{m_1}\}$ for $m_1 \in [M_1]$.

[0048] All clients share the supervised learning parameters θ_{m_2} for $m_2 \in [M_2]$.

[0049] Each client c keeps its own personalized learner weights $\pi_c(m_1, m_2)$, which satisfies $\sum_{m_1, m_2} \pi_c(m_1, m_2) = 1$.

[0050] Under the definition of the model, the optimization target will be equation 7 (M_1 or M_2 can be omitted when clear), as follows:

[0051] Equation (7):

$$\forall c \in \mathcal{C}, \max_{\pi_c, \theta} \mathbb{E}_{(x, y) \sim \mathbb{P}_c} \left[\log \left(\sum_{m_1, m_2} \pi_c(m_1, m_2) \mathcal{N}(x; \mu_{m_1}, \Sigma_{m_1}) P_{\theta_{m_2}}(y | x) \right) \right]$$

[0052] In a following step, the EM (Expectation-Maximization) algorithm is derived.

The EM (Expectation-Maximization) algorithm is used in machine learning to obtain maximum likelihood estimates of variables.

[0053] To reduce notation clutter, $m = (m_1, m_2)$ and $\Theta_m = \{\mu_{m_1}, \Sigma_{m_1}, \theta_{m_2}\}$. The model is then denoted as $P_{\pi_c, \Theta}(x, y) = \sum_m \pi_c(m) P_{\Theta_m}(x, y)$. Under this simplified notation, the EM algorithm is derived as follows:

[0054] First, $q_s(\cdot)$ is a probability distribution over $[M]$, where $s = (x, y)$. Also, for each sample, a first sample is drawn by the latent random variable $z \sim \pi_c(\cdot)$ and then sampling $(x, y) \sim P_{\Theta_z}(x, y)$. The EM algorithm is derived as follows:

$$\begin{aligned}
\sum_{x,y} \log (P_{\pi_c, \Theta}(x, y)) &\geq \sum_{x,y} \sum_{m \in [M]} q_s(m) \log \left(\frac{P_{\pi_c, \Theta}(z = m, x, y)}{q_s(m)} \right) \\
&= \sum_{x,y} \sum_{m \in [M]} q_s(m) \left[\log \left(\frac{P_{\pi_c, \Theta}(z = m)}{q_s(m)} \right) + \log (P_{\pi_c, \Theta}(x, y | z = m)) \right] \\
\text{Equation (8):} \quad &= \sum_{x,y} \sum_{m \in [M]} q_s(m) \left[\log \left(\frac{\pi_c(m)}{q_s(m)} \right) + \log (P_{\Theta_m}(x, y)) \right] \\
\text{Equation (9):} \quad &= \sum_{x,y} \sum_{m \in [M]} q_s(m) \left[\log \left(\frac{P_{\pi_c, \Theta}(z = m | x, y)}{q_s(m)} \right) + \log (\mathbb{P}_{\pi_c, \Theta}(x, y)) \right].
\end{aligned}$$

[0055] This evidence lower bound guides the EM algorithm:

[0056] For the expectation step (E-Step): Fix π_c and Θ , maximize Equation (9) via $q_s(m)$, and the optimal solution is equal to:

$$\begin{aligned}
q_s(m) &= P_{\pi_c, \Theta}(z = m | x, y) \\
&\propto \mathbb{P}_{\pi_c, \Theta}(z = m, x, y) = \pi_c(m) \mathbb{P}_{\Theta_m}(x, y).
\end{aligned}$$

[0057] For the maximization step (M-Step): Fix $q(\cdot | x, y)$, maximize Equation (8) via π_c and Θ , and the optimal solution will be:

$$\begin{aligned}
\pi_c(m) &= \frac{1}{N} \sum_{x,y} q_s(m), \\
\Theta_m &= \arg \max_{\Theta} \sum_{x,y} q_s(m) \log (\mathbb{P}_{\Theta}(x, y)).
\end{aligned}$$

[0058] Substitutions are then made for $m = (m_1, m_2)$ and $\Theta_m = \{\mu_{m_1}, \Sigma_{m_1}, \theta_{m_2}\}$.

Further, indexing is performed for the base component $\mathbb{P}_{m_1, m_2}(x, y) = \mathcal{N}(x; \mu_{m_1}, \Sigma_{m_1}) \cdot$

$\mathbb{P}_{\theta_{m_2}}(y | x)$, where $\mathcal{N}$ is a Gaussian distribution, and $\mathbb{P}_{\theta_{m_2}}(y | x)$ is represented by

some neural networks which outputs a distribution over the labels.

[0059] Under the aforementioned specific model, the updated rule can be rewritten as follows as:

[0060] For the expectation step (E-Step), the updated rules are rewritten as follows:

$$q_s(m_1, m_2) = \pi_c(m_1, m_2) \mathcal{N}(x; \mu_{m_1}, \Sigma_{m_1}) \cdot \mathbb{P}_{\theta_{m_2}}(y | x).$$

[0061] For the maximization step (M-Step), the updated rule can be rewritten as:

$$\begin{aligned} \pi_c(m_1, m_2) &= \frac{1}{N_c} \sum_{x,y} q_s(m_1, m_2), \\ \mu, \Sigma, \theta &= \arg \max_{\mu, \Sigma, \theta} \sum_{x,y} q_s(m_1, m_2) \log(\mathbb{P}_{\mu, \Sigma}(x)) + \sum_{x,y} q_s(m_1, m_2) \log(\mathbb{P}_{\theta}(y | x)). \end{aligned}$$

[0062] Further calculation will show it is equivalent to:

$$\begin{aligned} \pi_c(m_1, m_2) &= \frac{1}{N_c} \sum_{x,y} q_s(m_1, m_2), \\ \mu_{m_1} &= \frac{\sum_{x,y} \sum_{m_2} q_s(m_1, m_2) x}{\sum_{x,y} \sum_{m_2} q_s(m_1, m_2)}, \\ \Sigma_{m_1} &= \frac{\sum_{x,y} \sum_{m_2} q_s(m_1, m_2) (x - \mu^z)(x - \mu^z)^\top}{\sum_{x,y} \sum_{m_2} q_s(m_1, m_2)}, \\ \theta_{m_2} &= \arg \max_{\theta} \sum_{x,y} \sum_{m_1} q_s(m_1, m_2) \log(\mathbb{P}_{\theta}(y | x)). \end{aligned}$$

[0063] The complete algorithm derivation for Federated Learning with a Gaussian Mixture Model is as follows:

Algorithm for Federated Learning with GMM

1: for $t = 1, 2, \dots$ do

2: server broadcasts $\{\mu_m^{(t-1)}, \Sigma_m^{(t-1)}, \theta_m^{(t-1)}\}$ to all clients

3: for client $c \in [C]$ do

4: for component $m_1 \in [M_1], m_2 \in [M_2]$ do

5: for sample $s_{c,i} = (x_{c,i}, y_{c,i}), i \in [N_c]$ do

6: Set $q_{s_{c,i}}^{(t)}(m_1, m_2) \propto \pi_c^{(t-1)}(m_1, m_2) \cdot \mathcal{N}(x_{c,i}; \mu_{m_1}^{(t-1)}, \Sigma_{m_1}^{(t-1)}) \cdot \exp\left(-L_{\text{CE}}(s_{c,i}; \theta_{m_2}^{(t-1)})\right)$

7: end for

8: Set for all $m_1 \in [M_1], m_2 \in [M_2]$:

$$\pi_c^{(t)}(m_1, m_2) = \frac{1}{N_c} \sum_{i \in [N_c]} q_{s_{c,i}}^{(t)}(m_1, m_2)$$

$$\mu_{m_1,c}^{(t)} = \frac{\sum_{i \in [N_c]} \sum_{m_2 \in [M_2]} q_{s_{c,i}}^{(t)}(m_1, m_2) x_{c,i}}{\sum_{i \in [N_c]} \sum_{m_2 \in [M_2]} q_{s_{c,i}}^{(t)}(m_1, m_2)}$$

$$\Sigma_{m_1,c}^{(t)} = \frac{\sum_{i \in [N_c]} \sum_{m_2 \in [M_2]} q_{s_{c,i}}^{(t)}(m_1, m_2) (x_{c,i} - \mu_{m_1,c}^{(t)}) (x_{c,i} - \mu_{m_1,c}^{(t)})^\top}{\sum_{i \in [N_c]} \sum_{m_2 \in [M_2]} q_{s_{c,i}}^{(t)}(m_1, m_2)}$$

$$\theta_{m_2,c}^{(t)} = \arg \min_{\theta} \sum_{i \in [N_c]} \sum_{m_1 \in [M_1]} q_{s_{c,i}}^{(t)}(m_1, m_2) L_{\text{CE}}(x_{c,i}, y_{c,i}; \theta)$$

9: end for

10. client c sends $\{\mu_{m_1}^{(t)}, c, \Sigma_{m_1}^{(t)}, c, \theta_{m_2}^{(t)}, c, \gamma_c^{(t)}(m_1, m_2) = \sum_{i \in [N_c]} q_{s_{c,i}}^{(t)}(m_1, m_2)\}$

to the server

11. end for

12. for Gaussian component $m_1 \in [M_1]$ do

13. server aggregates

$$\mu_{m_1}^{(t)} = \frac{\sum_{c \in [C]} \sum_{m_2 \in [M_2]} \gamma_c^{(t)}(m_1, m_2) \mu_{m_1,c}^{(t)}}{\sum_{c \in [C]} \sum_{m_2 \in [M_2]} \gamma_c^{(t)}(m_1, m_2)}$$

$$\Sigma_{m_1}^{(t)} = \frac{\sum_{c \in [C]} \sum_{m_2 \in [M_2]} \gamma_c^{(t)}(m_1, m_2) \Sigma_{m_1, c}^{(t)}}{\sum_{c \in [C]} \sum_{m_2 \in [M_2]} \gamma_c^{(t)}(m_1, m_2)}$$

14. end for

15. for Supervised component $m_2 \in [M_2]$ do

16. Server aggregates

$$\theta_{m_2}^{(t)} = \frac{\sum_{c \in [C]} \sum_{m_1 \in [M_1]} \gamma_c^{(t)}(m_1, m_2) \theta_{m_2, c}^{(t)}}{\sum_{c \in [C]} \sum_{m_1 \in [M_1]} \gamma_c^{(t)}(m_1, m_2)}$$

17. end for

18. end for

[0064] FIG. 2 is a flow chart illustrating one embodiment of a method that uses a Federated Gaussian Mixture Model (Fed-GMM) to jointly model the joint probability of samples in each client 5, as illustrated in FIG. 1. Each client will keep its own data locally, however the clients 5 will communicate to a centralized server 7, e.g., computing server, supervised predictive components parameters and GMM parameters.

[0065] The method may begin with block 15. Block 15 can include input data being loaded into the edge devices. An edge device is any piece of hardware that controls data flow at the boundary between two networks. Edge devices fulfill a variety of roles, depending on what type of device they are, but they essentially serve as network entry -- or exit -- points. In the embodiment depicted in FIG. 1, the edge device may be a mobile computer, or it may be devices that collect data, e.g., via a motor vehicle or a specific institution, e.g., a banking institution or hospital. The edge devices can collect data from different regions, which may be different geographic regions, such as local data region 1, local data region 2, local data region 3, and local data region 4, etc.

[0066] In the example, of preparing a model that will predict likelihood of a purchase by a customer, e.g., the purchase of a vehicle, the data being input into the edge devices could be specifications on the product being sold. For example, if the product is a vehicle, the type of vehicle, e.g., coupe, sedan, truck, may be one type of data, as well as the color, make, model and price of the vehicle. Other data can include the region in which the sale is being made, and preferences by customers specific to that region.

[0067] Referring to block 20, following the input data being entered into the edge devices, the edge devices can train the local model 6. Referring to FIG. 3, training the local model 6 can include updating the Gaussiam Mixture Model components of the local model 6 at block 21. This is one mechanism by which heterogeneity in clients can be considered by the methods of the present disclosure. Training the local mode 6l can also include optimizing the model supervised component parameters at block 22. Training the local model can also include applying dimension reduction tools to high dimension data at block 23. For example, if the data is high-dimensional, dimension reduction tools are applied such as an auto-encoder.

[0068] The methods of the present disclosure utilize data from all the clients 5. Further the predictive models provided by the present method are suitable for all the different clients, e.g., the clients 5 from local data-region 1, local data-region 2, local data-region 3 and local data region 4.

[0069] Referring back to FIG. 2, the method may continue at block 25, which includes sending the local model parameters 6 to the server. Referring to FIG. 1, each of the clients 5 is in communication with a central server 7. The centralized server 7 may be in

communication to the clients wirelessly, and the centralized server 7 can be a cloud based system. Each client will keep its own data locally, however the clients will communicate to a centralized server, e.g., computing server, supervised predictive components parameters and GMM parameters. This provides a level of data security to the clients, as the data is not sent to the centralized server 7. Referring to FIG. 4, sending the local model parameters 6 to the server 7 may include sending the Gaussian Mixture Model parameter includes means variances together with supervised prediction mode parameters to the server 6. In a following step, the liner combination weights are also sent to the server 7.

[0070] Referring back to FIG. 2, in a following step the method may continue to block 30. Block 30 includes the server 7 aggregating the Gaussian Mixture Model parameters and averages the supervised components from the updated local models. Referring to FIG. 5, aggregation at the server 7 may include aggregating the Gaussian Mixture Model parameters including the means and variances at block 36; and aggregating the different conditional prediction component parameters from the local models at block 37.

Referring to block 1, the aggregated data is forwarded to a global model generator 11, at which new learned models are generated from the aggravated data. In one embodiment, aggregating the Gaussian Mixture Model parameters including the means and variances at block 36 can includes calculations in accordance with the algorithms from lines 12-14 of the complete algorithm derivation for Federated Learning with a Gaussian Mixture Model. In one embodiment, aggregating the different conditional prediction component parameters from the local models at block 37 can includes calculations in accordance

with the algorithms from lines 15-17 of the complete algorithm derivation for Federated Learning with a Gaussian Mixture Model.

[0071] Referring back to FIG. 2, in a following process step, at block 35, the aggregated global parameters are sent back to the local clients 5. The updated models, i.e., updated global model, is transmitted from the global model generator 11 to the central server 7, and ultimately to the clients 5. The local client 5 can then use the new learned model from the global model generator to update both the Gaussian Mixture Model and the predictive components.

[0072] Referring back to FIG. 2, in a following step, a determination is made whether convergence has been reached at decision block 40. A machine learning model reaches convergence when it achieves a state during training in which loss settles to within an error range around the final value. In other words, a model converges when additional training will not improve the model. If convergence has not been reached at decision block 40, the method can loop back to block 20, and repeat the steps associated with blocks 20, 25, 30 and 35, wherein another check for convergence is performed at block 40 following block 35. The steps associated with blocks 20, 25, 30 and 35 are looped until convergence is reached.

[0073] In some embodiments, when convergence is reached at decision block 45, the edge (client) devices have been fully trained to make predictions, and data relevant to predictions at the local edge devices may be entered into the edge devices to make predictions using a local model updated using the updated global model created using blocks 20, 25, 30 and 35. The trained model can then make a prediction at block 45.

[0074] In some embodiments, when there comes a new client with its own data, the model can calculate the probability that the new client data obey the mixture of distribution of existing training data.

[0075] FIG. 6 is an exemplary processing system for personalized federated learning under a mixture of joint distribution, in accordance with embodiments of the present invention.

[0076] The processing system includes at least one processor (CPU) 904 operatively coupled to other components via a system bus 902. A GPU 905, a cache 906, a Read Only Memory (ROM) 908, a Random Access Memory (RAM) 910, an input/output (I/O) adapter 920, a network adapter 930, a user interface adapter 940, and a display adapter 950, are operatively coupled to the system bus 902. Additionally, element 900 is a local model trainer following FIGs. 1 and 2. The local model trainer 900 can perform the steps described in block 35 of FIG. 2. More particularly, the local trainer 900 employs the global model that is generated by the global model generator 11, as depicted in FIG. 1. The global model generator 11 generally receives aggravated data from the central server 7, and generated a new global model employing the steps described in blocks 25 and 30 of FIG. 2.

[0077] A storage device 922 is operatively coupled to system bus 902 by the I/O adapter 920. The storage device 922 can be any of a disk storage device (e.g., a magnetic or optical disk storage device), a solid-state magnetic device, and so forth.

[0078] A transceiver 932 is operatively coupled to system bus 902 by network adapter 930.

[0079] User input devices 942 are operatively coupled to system bus 902 by user interface adapter 940. The user input devices 942 can be any of a keyboard, a mouse, a keypad, an image capture device, a motion sensing device, a microphone, a device incorporating the functionality of at least two of the preceding devices, and so forth. Of course, other types of input devices can also be used, while maintaining the spirit of the present invention. The user input devices 942 can be the same type of user input device or different types of user input devices. The user input devices 942 are used to input and output information to and from the processing system.

[0080] A display device 952 is operatively coupled to system bus 902 by display adapter 950.

[0081] Of course, the processing system may also include other elements (not shown), as readily contemplated by one of skill in the art, as well as omit certain elements. For example, various other input devices and/or output devices can be included in the system, depending upon the particular implementation of the same, as readily understood by one of ordinary skill in the art. For example, various types of wireless and/or wired input and/or output devices can be used. Moreover, additional processors, controllers, memories, and so forth, in various configurations can also be utilized as readily appreciated by one of ordinary skill in the art. These and other variations of the processing system are readily contemplated by one of ordinary skill in the art given the teachings of the present invention provided herein.

[0082] As used herein, the terms “data,” “content,” “information” and similar terms can be used interchangeably to refer to data capable of being captured, transmitted,

received, displayed and/or stored in accordance with various example embodiments.

Thus, use of any such terms should not be taken to limit the spirit and scope of the disclosure. Further, where a computing device is described herein to receive data from another computing device, the data can be received directly from the another computing device or can be received indirectly via one or more intermediary computing devices, such as, for example, one or more servers, relays, routers, network access points, base stations, and/or the like. Similarly, where a computing device is described herein to send data to another computing device, the data can be sent directly to the another computing device or can be sent indirectly via one or more intermediary computing devices, such as, for example, one or more servers, relays, routers, network access points, base stations, and/or the like.

[0083] As will be appreciated by one skilled in the art, aspects of the present invention may be embodied as a system, method or computer program product. Accordingly, aspects of the present invention may take the form of an entirely hardware embodiment, an entirely software embodiment (including firmware, resident software, micro-code, etc.) or an embodiment combining software and hardware aspects that may all generally be referred to herein as a “circuit,” “module,” “calculator,” “device,” or “system.” Furthermore, aspects of the present invention may take the form of a computer program product embodied in one or more computer readable medium(s) having computer readable program code embodied thereon.

[0084] Any combination of one or more computer readable medium(s) may be utilized. The computer readable medium may be a computer readable signal medium or a computer readable storage medium. A computer readable storage medium may be, for

example, but not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, or device, or any suitable combination of the foregoing. More specific examples (a non-exhaustive list) of the computer readable storage medium would include the following: an electrical connection having one or more wires, a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or Flash memory), an optical fiber, a portable compact disc read-only memory (CD-ROM), an optical data storage device, a magnetic data storage device, or any suitable combination of the foregoing. In the context of this document, a computer readable storage medium may be any tangible medium that can include, or store a program for use by or in connection with an instruction execution system, apparatus, or device.

[0085] Embodiments described herein may be entirely hardware, entirely software or including both hardware and software elements. In a preferred embodiment, the present invention is implemented in software, which includes but is not limited to firmware, resident software, microcode, etc.

[0086] Embodiments may include a computer program product accessible from a computer-usable or computer-readable medium providing program code for use by or in connection with a computer or any instruction execution system. A computer-usable or computer readable medium may include any apparatus that stores, communicates, propagates, or transports the program for use by or in connection with the instruction execution system, apparatus, or device. The medium can be magnetic, optical, electronic, electromagnetic, infrared, or semiconductor system (or apparatus or device) or a propagation medium. The medium may include a computer-readable storage medium

such as a semiconductor or solid state memory, magnetic tape, a removable computer diskette, a random access memory (RAM), a read-only memory (ROM), a rigid magnetic disk and an optical disk, etc.

[0087] Each computer program may be tangibly stored in a machine-readable storage media or device (e.g., program memory or magnetic disk) readable by a general or special purpose programmable computer, for configuring and controlling operation of a computer when the storage media or device is read by the computer to perform the procedures described herein. The inventive system may also be considered to be embodied in a computer-readable storage medium, configured with a computer program, where the storage medium so configured causes a computer to operate in a specific and predefined manner to perform the functions described herein.

[0088] A data processing system suitable for storing and/or executing program code may include at least one processor coupled directly or indirectly to memory elements through a system bus. The memory elements can include local memory employed during actual execution of the program code, bulk storage, and cache memories which provide temporary storage of at least some program code to reduce the number of times code is retrieved from bulk storage during execution. Input/output or I/O devices (including but not limited to keyboards, displays, pointing devices, etc.) may be coupled to the system either directly or through intervening I/O controllers.

[0089] Network adapters may also be coupled to the system to enable the data processing system to become coupled to other data processing systems or remote printers or storage devices through intervening private or public networks. Modems, cable

modem and Ethernet cards are just a few of the currently available types of network adapters.

[0090] As employed herein, the term “hardware processor subsystem” or “hardware processor” can refer to a processor, memory, software or combinations thereof that cooperate to perform one or more specific tasks. In useful embodiments, the hardware processor subsystem can include one or more data processing elements (e.g., logic circuits, processing circuits, instruction execution devices, etc.). The one or more data processing elements can be included in a central processing unit, a graphics processing unit, and/or a separate processor- or computing element-based controller (e.g., logic gates, etc.). The hardware processor subsystem can include one or more on-board memories (e.g., caches, dedicated memory arrays, read only memory, etc.). In some embodiments, the hardware processor subsystem can include one or more memories that can be on or off board or that can be dedicated for use by the hardware processor subsystem (e.g., ROM, RAM, basic input/output system (BIOS), etc.).

[0091] In some embodiments, the hardware processor subsystem can include and execute one or more software elements. The one or more software elements can include an operating system and/or one or more applications and/or specific code to achieve a specified result.

[0092] In other embodiments, the hardware processor subsystem can include dedicated, specialized circuitry that performs one or more electronic processing functions to achieve a specified result. Such circuitry can include one or more application-specific

integrated circuits (ASICs), field-programmable gate arrays (FPGAs), and/or programmable logic arrays (PLAs).

[0093] Reference in the specification to “one embodiment” or “an embodiment” of the present invention, as well as other variations thereof, means that a particular feature, structure, characteristic, and so forth described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, the appearances of the phrase “in one embodiment” or “in an embodiment”, as well any other variations, appearing in various places throughout the specification are not necessarily all referring to the same embodiment. However, it is to be appreciated that features of one or more embodiments can be combined given the teachings of the present invention provided herein.

[0094] It is to be appreciated that the use of any of the following “/”, “and/or”, and “at least one of”, for example, in the cases of “A/B”, “A and/or B” and “at least one of A and B”, is intended to encompass the selection of the first listed option (A) only, or the selection of the second listed option (B) only, or the selection of both options (A and B). As a further example, in the cases of “A, B, and/or C” and “at least one of A, B, and C”, such phrasing is intended to encompass the selection of the first listed option (A) only, or the selection of the second listed option (B) only, or the selection of the third listed option (C) only, or the selection of the first and the second listed options (A and B) only, or the selection of the first and third listed options (A and C) only, or the selection of the second and third listed options (B and C) only, or the selection of all three options (A and B and C). This may be extended for as many items listed.

[0095] The foregoing is to be understood as being in every respect illustrative and exemplary, but not restrictive, and the scope of the invention disclosed herein is not to be determined from the Detailed Description, but rather from the claims as interpreted according to the full breadth permitted by the patent laws. It is to be understood that the embodiments shown and described herein are only illustrative of the present invention and that those skilled in the art may implement various modifications without departing from the scope and spirit of the invention. Those skilled in the art could implement various other feature combinations without departing from the scope and spirit of the invention. Having thus described aspects of the invention, with the details and particularity required by the patent laws, what is claimed and desired protected by Letters Patent is set forth in the appended claims.

WHAT IS CLAIMED IS:

1. A computer implemented method for personalized federated learning comprising:
receiving at a central server (7) local models (6) from a plurality of clients (5);
aggregating a heterogeneous data distribution extracted from the local models (6);
processing the data distribution as a linear mixture of joint distributions to provide
a global learning model (10); and
transmitting the global learning model to the clients (5), wherein the global
learning model (10) is used to update the local model (6).
2. The computer implemented method of claim 1, wherein the receiving local
models (5) at the central server (7) includes sending parameters from the local models.
3. The computer implemented method of claim 1, wherein the receiving local models
(5) at the central server (7) does not include private data from the plurality of clients (5).
4. The computer implemented method of claim 1, wherein the local model (6)
updated by the global learning model (10) is used to predict an outcome from input data
applied to the local model updated by the global learning model (10).
5. The computer implemented method of claim 1, wherein an outcome predicted is
used to perform a sale at a price and specification in accordance with input data that
results in prediction of a successful sale.
6. The computer implemented method of claim 1, wherein training of the local
models (6) uses a log-likelihood maximization as training criterion.

7. The computer implemented method of claim 1, wherein the global learning model is a Federated Gaussian Mixture Model (Fed-GMM) that jointly models joint probability of samples in each client (5) of the plurality of clients (5).

8. The computer implemented method of claim 7, wherein the Fed-GMM that provides the linear mixture of joint distributions, further includes weight for parameters of the model that is personalized for each client (5).

9. A system for personalized federated learning comprising:

a hardware processor; and

a memory that stores a computer program product, the computer program product when executed by the hardware processor, causes the hardware processor to:

receive, using the hardware processor, at a central server (7) local models (6) from a plurality of clients (5);

aggregate, using the hardware processor, a heterogeneous data distribution extracted from the local models (6);

process, using the hardware processor, the data distribution as a linear mixture of joint distributions to provide a global learning model (10); and

transmit, using the hardware processor, the global learning model to the clients (5), wherein the global learning model is used to update the local model (6).

10. The system of claim 9, wherein the receiving local models (6) at the central server (7) includes sending parameters from the local models (6).

11. The system of claim 9, wherein the receiving local models (6) at the central server (7) does not include private data from the plurality of clients (5).

12. The system of claim 9, wherein the local model (6) updated by the global learning model (10) is used to predict an outcome from input data applied to the local model (6) updated by the global learning model (10).

13. The system of claim 9, wherein an outcome predicted is used to perform a sale at a price and specification in accordance with the input data that results in prediction of a successful sale.

14. The system of claim 9, wherein training of the local models uses a log-likelihood maximization as training criterion.

15. The system of claim 9, wherein the global learning model (10) is a Federated Gaussian Mixture Model (Fed-GMM) that jointly models joint probability of samples in each client of the plurality of clients.

16. A computer program product for personalized federated learning, the computer program product can include a computer readable storage medium having computer readable program code embodied therewith, the program instructions executable by a hardware processor to cause the hardware processor to:

receive, using the hardware processor, at a central server (7) local models (10) from a plurality of clients (5);

aggregate, using the hardware processor, a heterogeneous data distribution extracted from the local models (10);

process, using the hardware processor, the data distribution as a linear mixture of joint distributions to provide a global learning model (10); and

transmit, using the hardware processor, the global learning model (10) to the clients (5), wherein the global learning model (10) is used to update the local model (6).

17. The computer program product of claim 16, wherein the receiving local models (6) at the central server (7) includes sending the parameters from the local models.

18. The computer program product of claim 16, wherein the receiving local models (6) at the central server (7) does not include private data from the plurality of clients (5).

19. The computer program product of claim 16, wherein the local model (6) updated by the global learning model (10) is used to predict an outcome from input data applied to the local model updated by the global learning model (10).

20. The computer program product of claim 16, wherein an outcome predicted is used to perform a sale at a price and specification in accordance with input data that results in prediction of a successful sale.

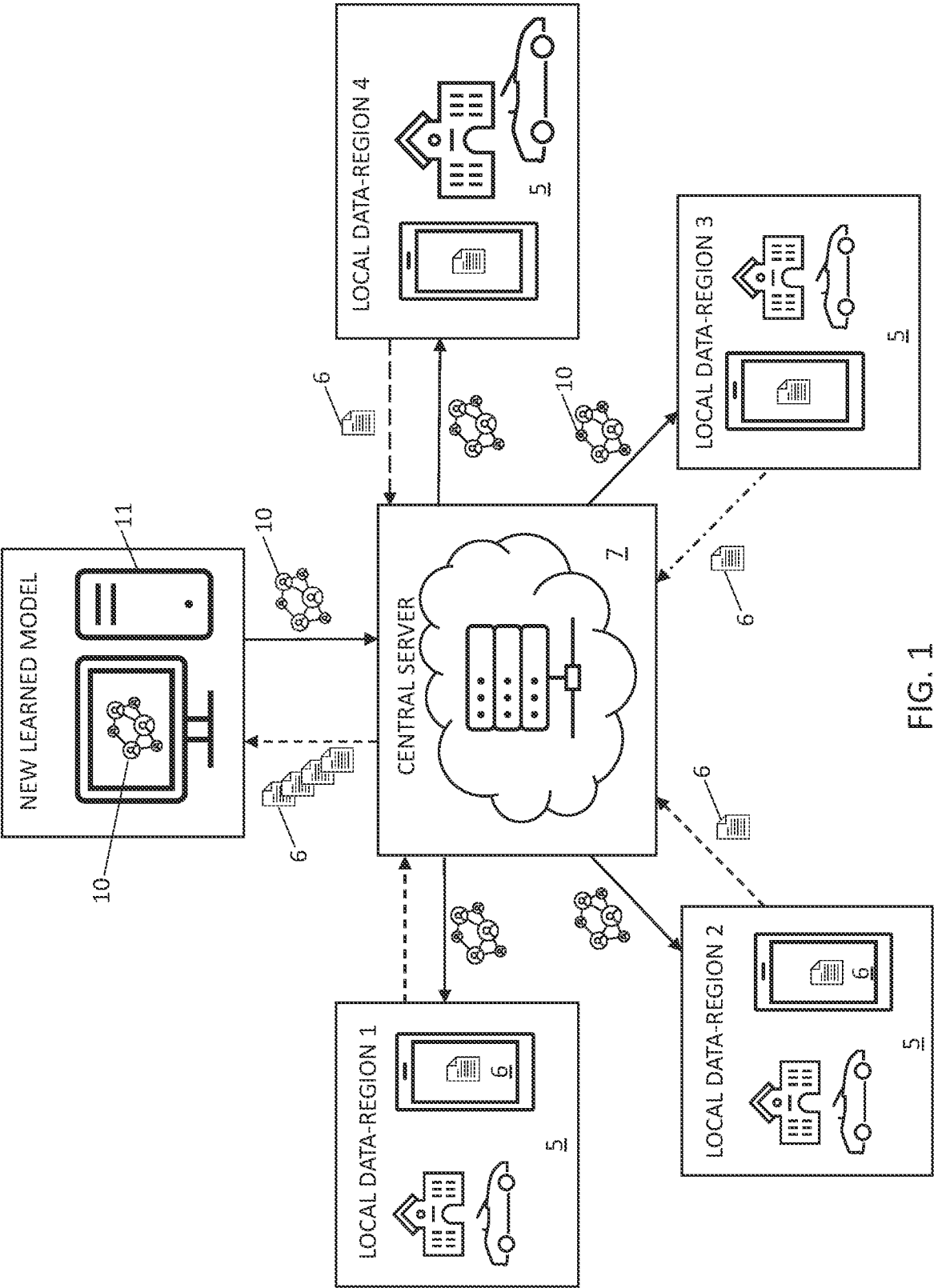


FIG. 1

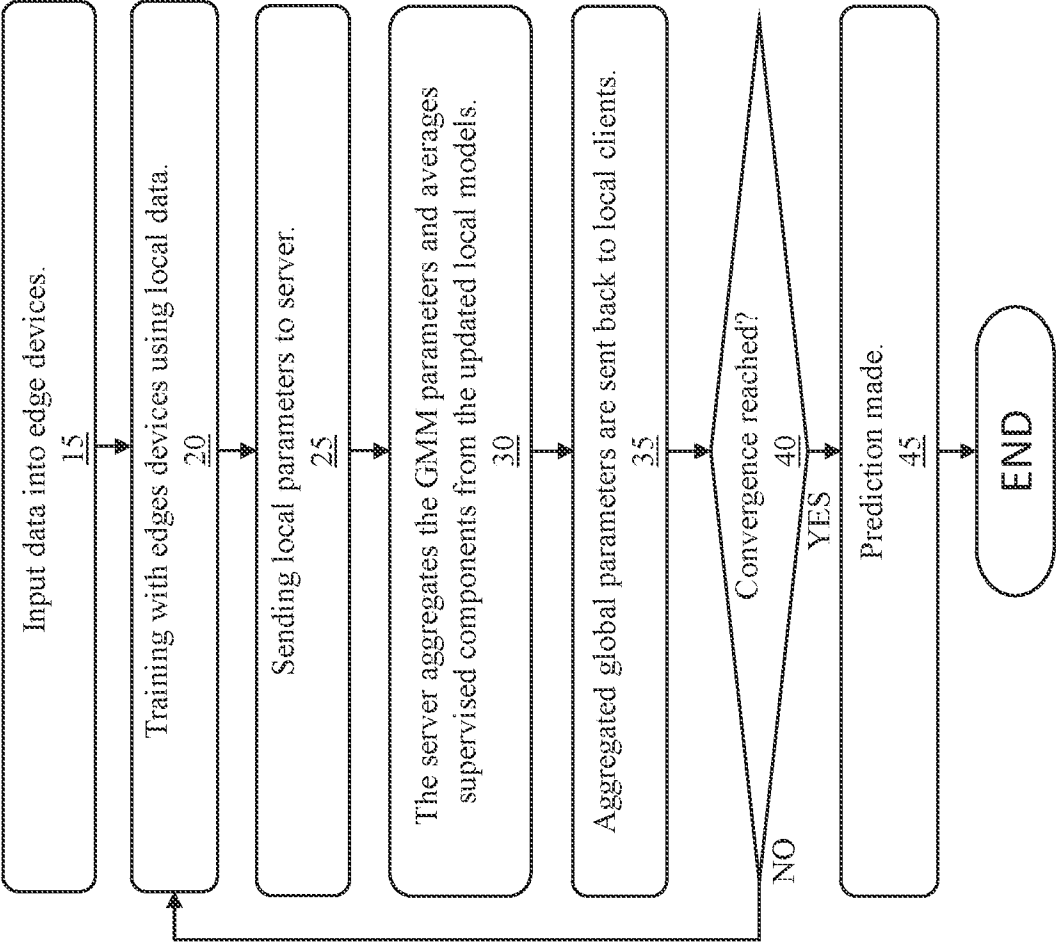


FIG. 2

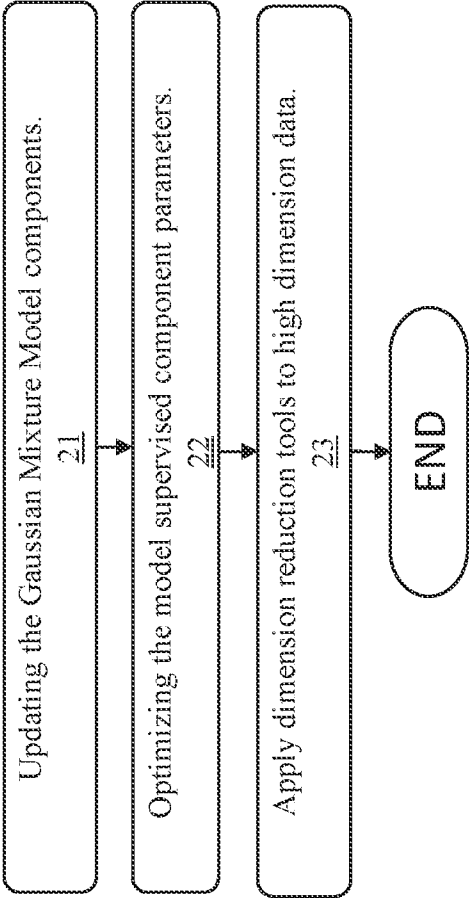


FIG. 3

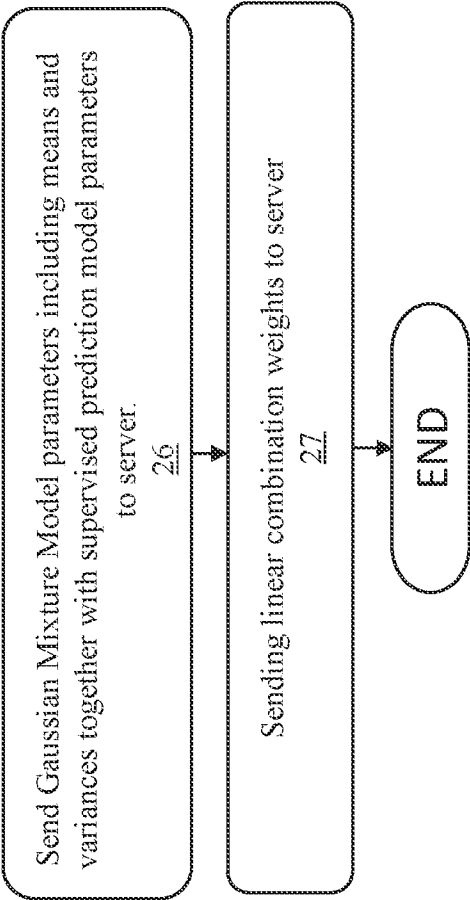


FIG. 4

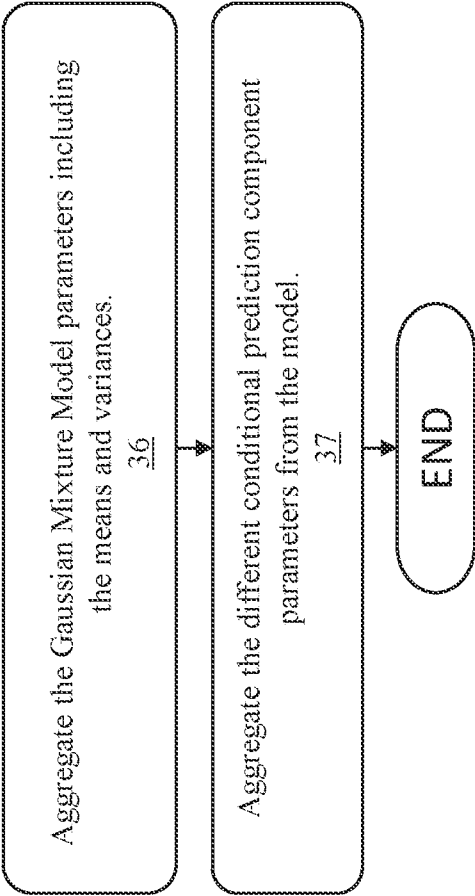


FIG. 5

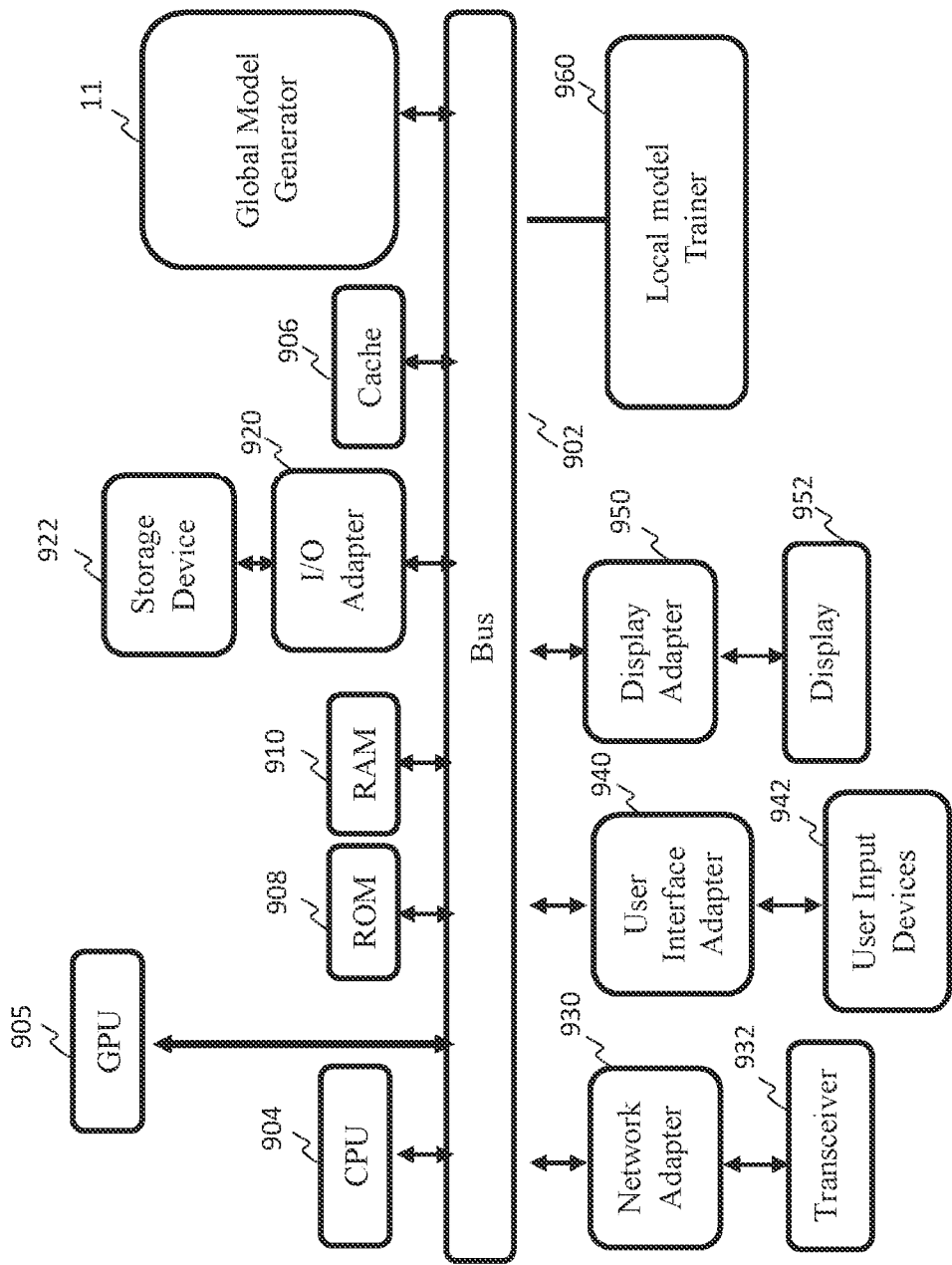


FIG. 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2023/032687

A. CLASSIFICATION OF SUBJECT MATTER**G06N 20/20**(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06N 20/20(2019.01); G06N 20/00(2019.01); G06N 3/02(2006.01); G06N 3/04(2006.01); G06N 3/08(2006.01);
H04L 12/24(2006.01); H04L 41/0823(2022.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS(KIPO internal) & Keywords: federated learning, receiving, parameters, aggregating, linear mixture, global model,
local model**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2021-0065002 A1 (FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V.) 04 March 2021 (2021-03-04) paragraphs [0049]-[0050], [0073], [0166]; and figures 3, 4c	1-20
Y	US 2022-0114475 A1 (RUI ZHU et al.) 14 April 2022 (2022-04-14) paragraphs [0017], [0046]-[0047], [0097]	1-20
A	WO 2021-121585 A1 (TELEFONAKTIEBOLAGET LM ERICSSON (PUBL.)) 24 June 2021 (2021-06-24) paragraphs [0112]-[0117]; and figures 12-14	1-20
A	WO 2022-171354 A1 (NOKIA TECHNOLOGIES OY) 18 August 2022 (2022-08-18) paragraphs [0062]-[0069]; and figures 9-10	1-20
A	US 11443245 B1 (ALIPAY LABS (SINGAPORE) PTE. LTD.) 13 September 2022 (2022-09-13) column 8, line 40 – column 10, line 40; and figure 3	1-20

☐ Further documents are listed in the continuation of Box C.
☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

05 January 2024

Date of mailing of the international search report

05 January 2024

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
189 Cheongsu-ro, Seo-gu, Daejeon
35208, Republic of Korea

Authorized officer

YANG, JEONG ROKFacsimile No. **+82-42-481-8578**Telephone No. **+82-42-481-5709**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/US2023/032687

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2021-0065002	A1	04 March 2021	CN	112424797	A	26 February 2021
				EP	3794515	A1	24 March 2021
				WO	2019-219846	A1	21 November 2019
				WO	2019-219846	A9	04 March 2021
US	2022-0114475	A1	14 April 2022	CN	116745780	A	12 September 2023
				WO	2022-073320	A1	14 April 2022
WO	2021-121585	A1	24 June 2021	EP	4078899	A1	26 October 2022
				US	2023-0010095	A1	12 January 2023
				US	2023-0019669	A1	19 January 2023
				WO	2021-123139	A1	24 June 2021
WO	2022-171354	A1	18 August 2022	EP	4292255	A1	20 December 2023
US	11443245	B1	13 September 2022	None			