Continuous Federated Learning of Global Policies in Coalition Environments

Yaniv Aspis*, Daniel Cunnington*†, Mark Law*, Alessandra Russo*,
Krysia Broda*, Jorge Lobo*, Ankush Singla‡, Elisa Bertino‡ and Dinesh Verma§
*Imperial College London, London, UK
†IBM Research Europe, Winchester, UK
‡Purdue University, West Lafayette, IN, USA
§IBM Research, Yorktown Heights, NY, USA

***Abstract*—In coalition environments, multiple parties may need to collaboratively solve a coalition-based (or global) decision-making task in a federated manner, by using models trained autonomously from their own local datasets, and sharing as little information as possible. Instead of adopting a centralised federated learning approach, which may suffer from a single central server bottleneck, we propose a *global policy learner* framework that learns global policies by taking into account just the predicted outcomes of the locally trained models, hence requiring minimal information sharing among the local parties. A global decision-making task on new unlabelled data is solved by (i) collecting local predictions from the parties, through their own learned models, together with associated level of confidence, (ii) combining them to produce a label for the data, and (iii) using the symbolic learner FastLAS to learn from this labelled data a global policy capable of performing global predictions. Such policies can be continuously updated as new data become available and/or local models get retrained when needed. The learned global policies can be used in situations where the communication with the parties is unreliable due to failures or out-of-range drop out.**

I. Introduction

In a coalition setting, multiple parties are required to collaboratively solve global decision-making tasks by using models that are autonomously learned from their own local datasets, and sharing as little information as possible. Coalitions are very dynamic, with parties joining or leaving at any point, operating with unreliable communication due to failures or drop out in highly dynamic environments. Local trained models are therefore also subject to change. To accommodate decision making in such an environment, a novel framework is needed that differs from conventional federated learning in that global decision-making policies can be efficiently learned from predicted outcomes of local models. To be of practical use, a global decision-making policy needs to be as dynamic as its environment. It needs to (i) use predicted information available from the local parties whilst also not being dependent on their availability, (ii) accommodate a wide variety of models that local parties may choose to train, and (iii) require only minimal information to be shared by the parties.

Current approaches for federated learning come short for some of these requirements. For instance, they train a single model by sharing it across multiple nodes, each contributing based on their own local datasets [1]. As such they strongly rely upon the communication between parties and need to adopt the same machine learning technique. In an ensemble approach, on the other hand, different models can be learned at the different local nodes, but with respect to a fixed common dataset [2] which has to be accessible by each node. Neither of these two techniques is suitable for the coalition setting described above.

In this paper, we propose a novel framework for learning global policies for coalition-based decision-making, which properly satisfies the requirements highlighted above. Global policies are learned by taking into account predictions and associated confidence level of each of the different party's models. These form a continuously evolving global dataset that is used by a symbolic learning system, called FastLAS [3], to learn interpretable global policies. These policies can be dynamically updated to adapt to changes in the coalition and to support decision making on new data points, even when (some of) the local parties may not be reachable.

II. Framework

A. Definitions

A coalition is assumed to be formed by a set of *local agents* $\mathcal{AG}_1, \mathcal{AG}_2, ..., \mathcal{AG}_N$ (which may change over time) and a *global policy agent* $\mathcal{AG}_p$. Agents are assumed to solve a similar classification task, but using their own local dataset, which they do not wish to share. Let $\mathcal{X}$ denote a universal set of raw data points, $\mathcal{F}$ a universal set of features and $\mathcal{O}$ a universal set of labels. Different data in $\mathcal{X}$ may be characterised by different sets of features in $\mathcal{F}$. Specifically, each agent $\mathcal{AG}_i$ has its own subset of features $F_i \subseteq \mathcal{F}$, related to its local raw data. Local agents do not need to share the same set of features and may have their own subset of labels $O_i \subseteq \mathcal{O}$. Local agents do not need to share labels with each other, but the global agent is aware of the labels each agent uses to make predictions.

Each local agent has its own dataset $\mathcal{D}_i$ of labelled examples. An example is a triple (x, o, w) where $x \in \mathcal{X}$, $o \in \mathcal{O}$ and $w \in \mathbb{R}$ is a *weight*. The weight indicates the example's level of trust. Intuitively, the example indicates that a data point x should be classified with label o with level of trust w. i.e. not predicting the label o causes the model to pay a penalty w. A negative weight indicates that it is preferable for the agent not to predict the label o for the data point x.

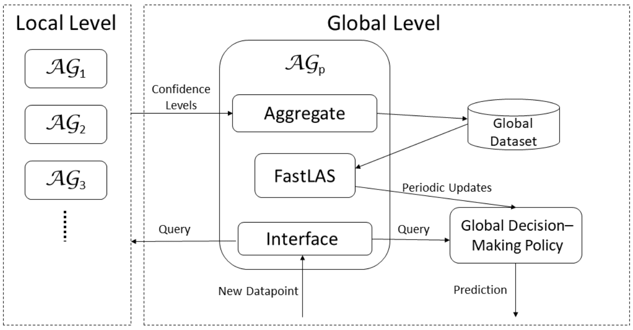

Fig. 1. Framework architecture. New incoming data points are evaluated using the current global policy and a prediction is returned. In parallel, they are sent to available local agents for confidence values. These values are aggregated, then a new example is constructed and stored in a global dataset. FastLAS is periodically invoked to learn and update the global policy.

Using the local dataset $\mathcal{D}_i$, an agent $\mathcal{AG}_i$ can train its local model M_i to solve the classification task. In principal, there are no restrictions on what this model may be. The only requirement is that given a data point x and label $o \in O_i$, the model is able to return a confidence level $\text{conf}_i(x, o)$ indicating how confident it is that x should be labeled o. For example, if M_i is a neural network that outputs posterior probabilities, confidence values can be estimated if the output is well-calibrated.

B. Global Policy

At the global level, a new unlabelled data point x is handled in three stages: Predictions about the data point are *collected* from available local agents. The agents are queried for their opinion on x and return the values of $\text{conf}_i(x, o)$ for each label o they are aware of. One can think of these values as forming a *confidence matrix* $C \in \mathbb{R}^{N\times|\mathcal{O}|}$ where $C_{i,o} = \text{conf}_i(x, o)$ if the i'th agent has a confidence value for o, otherwise $C_{i,o} = 0$.

Next, the various confidence values for a label o are aggregated to form an entry of a *confidence vector* $\vec{\text{conf}}(x)$ of size $|\mathcal{O}|$. This is done by means of an aggregation function $agg : \mathbb{R}^N \to \mathbb{R}$ that is applied across the confidence values of all the agents, i.e. $\vec{\text{conf}}(x)_o = agg\,(C_{1,o}, C_{2,o}, ..., C_{N,o})$. We thus have an aggregated confidence value per label.

For each label o, an example (x, o, w) is constructed with $w = \vec{\text{conf}}(x)_o$. These examples are added to a continuously growing *global dataset*. At first, this dataset would be small. However, over time it will grow to reflect the opinions of the local agents more accurately. A global policy learned from this dataset will thus take the local agents' beliefs into account. As the dataset grows, the policy can be updated to reflect the dynamic nature of the environment. To perform the learning itself, we use FastLAS, a state-of-the-art system for Inductive Logic Programming. The policies learned by FastLAS are guaranteed to be optimal with respect to coverage of the examples. They are also human interpretable, allowing for decisions made by the global agent to be explained.

Once a global policy is available, it can always be queried for predictions on new datapoints directly, without querying the local agents first. Doing so is often beneficial, as the process of querying agents and updating the policy can be slow or communication might not be available. The longer process of information collection, aggregation and updating should be done offline.

The nature of the aggregation function is critical to the success of this method. As a simple approach, one can take $\text{agg}\left(\vec{\text{conf}}\,(x)\right)_o = \sum_i C_{i,o}$. This approach however is likely to be non-optimal. Since confidence values can be provided by highly distinct models, they will generally be of different scale, and require normalisation first. Even after normalisation, a simple summation may not suffice. Certain models may be trusted more than others under certain situations. For instance, one model may be more reliable in predicting a certain label than another. Weighing the different confidence levels and taking a weighted average may prove more beneficial.

As mentioned, the policy should be updated periodically whenever sufficient new examples are available. Initiating an update for every new data point is likely to be inefficient, as a single data point is unlikely to change the policy much. One should instead apply updates in batches, for example whenever M new examples have been added to the global dataset, M being a tunable parameter. A more dynamic approach is to attempt to detect shifts in the data distribution by comparing the predictions of the global policy with those of the local agents. If the global policy often makes a different prediction from those of the local models, an update is likely to be in order.

III. Discussion

In this paper we lay the foundations of a framework for continuous learning of global policies in dynamic coalitions. A proof of concept experiment has been carried out on several datasets.

This work is related to previous work on policy-based ensembles in coalition environments [4], where local agents are queried for their opinion on new data points by means of a broadcast policy and results are aggregated by means of a combination policy. These policies are generated from hand-crafted rules, rather than learned. They require agents to share more information about their dataset or models, depending on the nature of the policy, and are about how to combine predictions. Predictions based on a broadcast and combination policies can at times be more accurate than those made by a learned global policy, as they take into account the data available to the local agents. But they require that at least some agents are available, otherwise a prediction cannot be made, and as they are hand-crafted, may not be optimal. In contrast, our framework allows the learning of new interpretable global decision making policies, rather than combination policies, which are guaranteed to be optimal with respect to the seen data. As they are learned, these policies are automatically adapted over time and used to provide global predictions, regardless of the availability of the agents. Combining the two approaches, for instance by using a more-informed combination policy as an aggregation function, may be a promising approach that we wish to explore in future work.

ACKNOWLEDGMENT

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

[1] J. Konečný, H. McMahan, D. Ramage, and P. Richtárik, "Federated optimization: Distributed machine learning for on-device intelligence," 10 2016.

[2] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1249, 2018.

[3] M. Law, A. Russo, E. Bertino, K. Broda, and J. Lobo, "Fastlas: Scalable inductive logic programming incorporating domain-specific optimisation criteria."

[4] D. C. Verma, E. Bertino, A. Russo, S. Calo, and A. Singla, "Policy-based ensembles for multi domain operations," in *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*, T. Pham, L. Solomon, and K. Rainey, Eds., vol. 11413, International Society for Optics and Photonics. SPIE, 2020, pp. 48 – 56. [Online]. Available: https://doi.org/10.1117/12.2558727