Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2021, 8(3):186-192**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Raspberry Pi Utilization in a Federated Machine Learning Domain

Barida Baah[1], Chioma Lizzy Nwagbo[2], Shedrack Mmeah[3]

[1]Department of Computer Science, Ebonyi State University, Abakaliki- Nigeria
[2]Department of Computer Science and Robotics Education, College of Education, Nsugbe, Anambra State, Nigeria
[3]Department of Computer Science, Ken-Saro Wiwa Polytechnic, Bori, Rivers State - Nigeria

Abstract Privacy and security is becoming a major challenge when it comes to the distributed systems like the federated machine learning system especially when data are been transmitted or learned on a network , this necessitated the reasons for this research work which is all about wireless federated machine learning process using a raspberry Pi. The raspberry Pi 4 is a single hardware board with built in Linux operating system. Data set of names where used from different languages and then we develop a training model using recurrent neural network to train this names compare to the names in the existing language like French, Scottish to predict if the names are from any of this language, this is done wirelessly with the Wi-Fi network in a federated machine learning environment for experimental setup with PySft's that is installed in the python environment. The system was able to predict the name from which the language it originated from, the methodology that is implore in the research work is the Rapid Application Development (RAD). The benefits of this system are to ensure privacy, reduces the computing power, and most importantly it is cost effective.

Keywords Utilization, Federated Machine learning, Raspberry, Pi, RNN

Introduction

Applied learning strategies include integrating training data into a specified database or database. For example, if the first e-commerce company wants to set up its customer information module for purchasing its products, it would run the information on the data collected from the application or website. The data may be related to the time spent on the actual product page Products that have been searched but not purchased and products that have been purchased. The data should be sent and received through an encoder or data source [1].

Although the comparisons could be quite simple, computer history could substitute for what Federated Learning is all about. There has been a huge gap in the early days of information technology, making the most advanced computer simulations. Finally, the researcher switched to a remote system where computers were distributed between client computers, clients, and internal servers [2].

The Federated Learning structure uses the same model. Machine learning machines are distributed over the systems of computer equipment, rather than on large, centralized systems. This computer model, while initially thought of, would not have been hand-made, as the mobile computer would be too slow to run any Machine Learning module.

Consequently, learning improved in the near-to-late 2018s. Since billions of downloads, equipped with AI chips and high-capacity computers, starting with the Samsung S9, or Apple X series, some of the machine learning (ML) models should run in that and focus on such phones, delivered in the next 3-5 years.

Review of Related Works

Heyam H. Al-Baity *et al* [3] proposed an effective web service discovery using feature selection, it aims is to employ feature selection in order to search for subset characteristics that can increase the accuracy classification and also enhanced the pace of the traditional classifiers using machine learning.

Laachemi *et al* [4] apply the use of support vector machine with the classification that improved on its accuracy level based on a stochastic local search (SLS),at the initial stage**,** the web service quality dataset attributes utilizes the support vector machine classifier to find solutions that is optimal as they have been scaled and its values resized. SLS-SVM approach has an accuracy which was 84.86%, that is far better than the SVM with similar classifier.

Liu *et al* [5] proposed the integrated of SVM classifier and the latent Dirichlet allocation (LDA) based models which is used big-scale services to classify and reduced the cost of labeling manual services for training. Algorithm of LDA was used to remove high-level topics from services. The base classifier for this technique is the SVM was used due to the facts that its does well on the classification text. Web services classifier is an important aspect based on its descriptions. Furthermore, with the introduction of a database-based learning strategy that is active was used to reduced manual labeling services cost that is needed to build and trained the set. Manually labeling of services was done in the preprocessing stage, with descriptions information based on its capabilities. Their research takes a looked at the groups with high numbers of services and then select 10 groups. These resulted to increase in the service number of the LDA-SVM which provides the results that are more accurate compare to the SVM. The efficacy of classification frame work in its active learning service was clearly shown in their experimental result that was obtained.

Mustafa *et al* [6] opined multi-layer perception neural network (MLPNN) novel classification model which applies optimization through search tabu**.** It utilizes multi-layer perception neural network model to inspire the neuroscience which was utilize in the prediction of process. MLPs were utilized back propagation in conjunction with Levenberg-Marquardt algorithm that trained the multi-layer perception classifier and TS in the optimization of the classifier. The quality of web-service dataset is like the dataset used [4]. The MLP-TS demonstrated better results of experiment.

Federated Machine Learning

It is a trained process in machine learning algorithm, taking the case of deep learning neural networks as an instance its works on replicated datasets that are local and has a node that local without explicitly substituting the various samples of data. Its process involves training the local models given a local data samples and substituting parameters of its nodes that is local to some occurrences in order to create global model that each and every node will be shared. In other words, the FL are rising distributed technique that is particularly aware of the difficulties which includes isolation and constraints of resources. Processing power of on-device utilizes to untapped personal data by performing the training procedure in any given distributed way and data kept where been created. It also provides easily accessibility through the process of federated learning ML which has expansion of a unique federated learning technique as it is proposed by Google recently [7-8].

The major dissimilarity between the FL and DL depend on the properties of the datasets that are local [9]. In a distributed learning environment, its original aim to perform parallelize power of computing in a FL where its originally aim is to train on different dataset. On the other hand, the distributed learning is to train a single procedure which is achieve via a multiple server, and its factor are assumption of datasets that are local and comparably decentralized and are of equal size.

Kunal Chandiramani [10] stated that different data quantity can be important to trained a model, thou they adhere to recent procedures that are confidential and localized data in order to stop organization from free utilization**.** Problems solving attempts are given for privacy-preserving device from model training for any given distributed data rather data are collected and sent to a server that is centralized, for each training devices model are able to send data separately. As a result, it has the would-be when substituting today foremost model of centralization of its computing.

Kunal Chandiramani [10] also proposed federated learning to model for a fashion MNIST dataset his result indicate that privacy maintenance in federated learning is not only job but it quick and allow operation at a scale even with low computing power and mobile devices that has been utilize. FL is a decentralized technique is built to provide an effective preserving-privacy machine learning environment for any decentralized system [11].

In FL, machine learning model is trained to create data at computer owners' ends s and then do the coordination of the server that also used to create model that is global and share the ML knowledge among the decentralized entities via devices. Due to fact that the actual data will not depart from the owners' data devices, FL provide privacy to raw data whenever, ML models indicate weakness to attacks from privacy inference in a way that the model will memorize these attacks and their inference membership, will be centered on accessing responsive data from its trained ML models even under a given black-box settings [12-13].

J. Konecny [14] proposed a training model for machine learning, traditional machine learning utilizes central technique for aggregated trained data on one machine. Alternatively, the technique for training centralized intrusive privacy, specifically, in the case when mobile devices gather data which contain information that is vital for owners. Transmitting all data gather for mobile devices that not feasible because of the limitations in communication resource. With such consideration, the federated learning (FL) concept will enable the large corps for trained distributed central data that are residing on mobile devices.

Raspberry Pi

Raspberry Pi is one-board computer just like Intel Galileo, and BeagleBone. It has a low-cost development device use for educational purposes testing. The Intel Galileo, BeagleBone and Raspberry Pi are fully programmable and customizable, and it has those features that are needed in the implementation small and low-cost Internet of Thing (IoT) devices [15].

The most popular of these three is the RaspberryPi according to the research field work keyword search. Computer one-board technology arising in the prototypes comes most often when prototype are developed it is a difficult experience and also costly in nature as a results of the costs of hardware design, software design and developing hardware manufacturing and building. When utilizing one-board computer, its prices reduces simply because of Hardware solutions that are ready-to-use already which exist which has embedded Linux ready-to-use software.

A twin study of [16] and [17] Introduced IoT-based E-learning which was developed testbed based on integration of 5 Raspberry PIs and a sensor microwave. They identify effects which are responsible for the testbed controlled as follows: Control chair Vibrator, Control Light, control Smell, Sound Control and a Remote-Control Socket, improvement and e-learners motivation is stimulated by utilization of this testbed. It utilizes the optimization of the link state Rout protocol technique in the testbed software.

Yamada, M. *et al* [16] proposed research are of the same authors which same issue were taking care of more profoundly. Oda M., *et al* [17] it test the bed network, it communication was tested and it results were shown.

Mahmoud and Qendri [18] also proposed a shielded sensor for **s**ensorial platform, which was used for Raspberry Pi. Their major purpose is transition process of Raspberry Pi into the platform of IoT, there research classified hardware into development process**.** It comprises of light sensors**,** accelerometer, temperature, pressure, and touch pad. The developed shield was done via crowded means of funding. It does not contain any functional tests. Regarding the firmware software testing with Raspbian operating system however the test case result was not mentioned.

Raikar, M. M. *et al* [19] deal on education in their research work. In their study it introduces the problems of Internet-of-Things while the later deals with Python programming language teaching**.** Their study Raspberry P is and sensor sets were utilized for learning purposes. They were able to build prototypes using the hardware. Their studies were not able to touch testing of the built prototypes. Their main aim was more on education than the other area of research work. The requirement specification of the system was however, included in their studies.

Maksimovic, Vujovic and Perisic [20] proposed internet-of-things based e-health systems they pointed out in their research the economic effects of the internet-of-thing applications, especially the aspect of healthcare applications economic growth. Their works compare various applications: sensor thee-health is shield V2.0 for

Raspberry Pi while others is a customized sensor made for measuring body system. The two applications perform data collection and sends it to back to the server. Their works do not contain any special testing part, but their work comprehensively points out the issues of security data collections.

Hentschel, Jacob, Singer and Chalmers [21] suggested system for campus smart that depends on Raspberry Pis**.** It utilizes hardware, the services of software and their architecture of the hardware which has a sensors and Raspberry PIs. Services of the cloud used to store data their works shows the technology of sensor-to-sensor and data movement delay-tolerant which is for data that are not urgent due to disruption of network connection. Several cases of system utilization have been described in their work like the free meeting room, the temperature of room, occupant of room, robotic support infrastructure and custom event triggering. But it does not give a description of the physical test cases. However, the enhancement of hardware design was brought in by replacement of the various sensor's types.

System Design

In this research paper work we will define all the technical and functional requirements. All the software and hardware specifications that are used must be setup in this phase. In starting this design phase, we get the initial overview of the processes of using the software. In this research work, Raspbian will be used as operating system for Raspberry Pi and coding will be done in python.

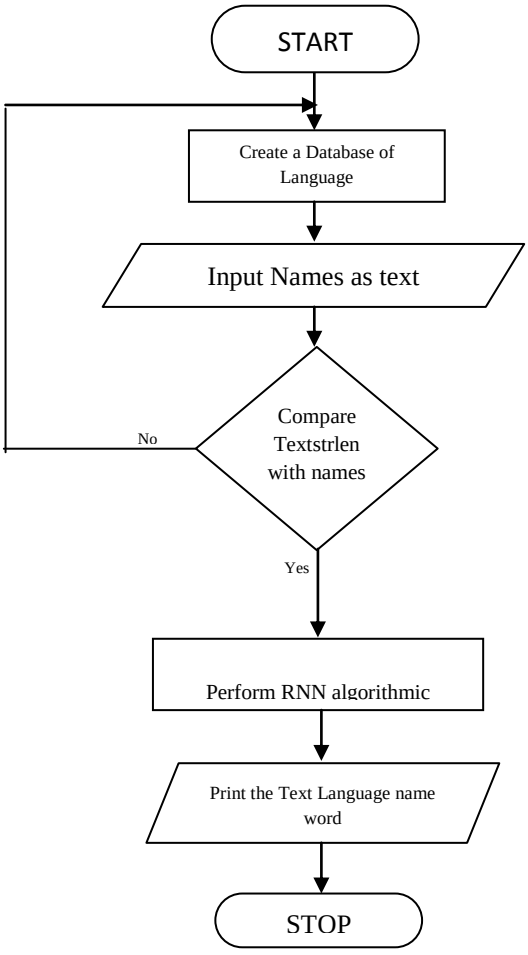

Figure 1: A Proposed Flowchart for the federated learning of text classification

Experiment Setup

In the experiment set up section, we present the setting of the experiment process for hardware and software process on collected datasets. The Federated Learning was first implemented on the laptop (coordinating server)

before using the raspberry pi. A Raspberry PI 4 running Raspbian 20.04 connected to the internet via SSH and Wi-Fi network with each PI having is configured with a static IP address configured on them for convenience of addressing. The static IP address are 192.168.0.77 and 192.168.0.78 respectively. Figure 2 show a laptop running on Ubuntu Linux connected to the same LAN with the raspberry Pi via Wi-Fi.

Figure 2: An experiment setup of the system

Figure 2 shows the successful experiment setup for the two raspberry Pi 4 installation process

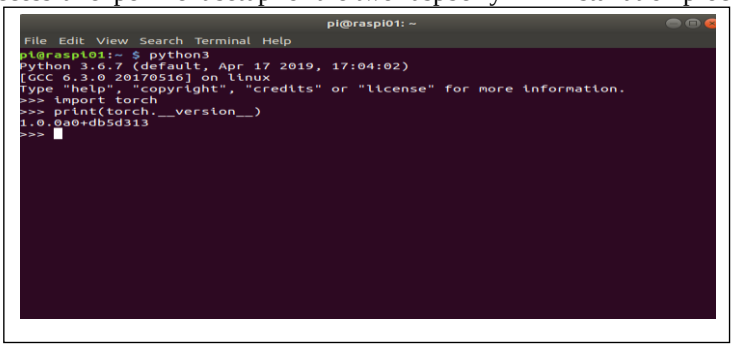

Figure 3: Successfully imported Pytorch in Python

Figure 3 above shows the successful installation of imported Pytorch in Python environment

Experiment 2: Training the Recurrent Neural Network on Raspberry Pi

The Recurrent Neural Network will be trained for classification of a person's surname to its most likely language of origin in a federated way, using workers sever running on the two Raspberry PIs that are already equipped with python3.6, PySyft, and Pytorch. A character-level Recurrent Neural Network treats words as a series of characters, outputting a prediction and "hidden state" for every character, feeding its previous state into each next step. We then take the final prediction, which is the output, i.e. which class the word belongs to. Hence the training process continuous sequentially in a character-by-character manner through the different hidden layers.

Result and Discussion

We plot the results using Matplotlib's pyplot. The plot will show us the learning rate of our recurrent neural network. The plot will show us the learning rate of our network.

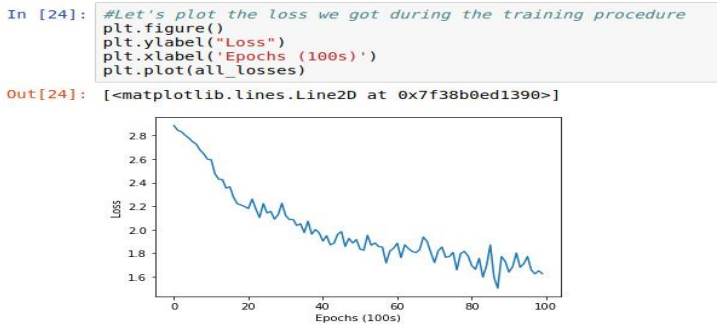

Figure 4: A Screenshot showing learning rate of the network

Testing the Trained RNN (Predicting names)

We can do inference locally by getting model back from Raspberry Pi and we can test using different models from each worker or e could also average all the model parameter from different worker, figure 21 shows testing on two(2) separate model from two(2) raspberry Pi("bob", "alice"). We defined a function that take in a name and return the likely languages base on our database list created.

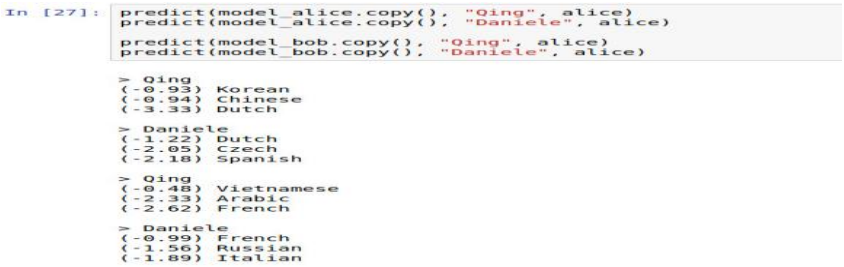

Figure 5: Screenshot returning likely languages of names

Conclusion

In conclusion, it is usually tasks demanding to perform classification of text using Recurrent Neural Network (RNN) algorithm with long-Short Time Memory Network(LSTM) in python due to the fact that the sequence of input text or name over spaces of time is sometime vary in length with a very large vocabulary of input symbols as a result its usually require a complex model to learn the long-term dependencies or context between symbols in its input sequences. Our develop system was able to achieve the following:

i. A platform for text name classification to predict name in language
ii. A platform that integrate the Raspberry Pi 4 with the preloaded Ubuntu Linux Operating System to ensure privacy of the user training model

References

[1]. F Rottenberg, T Choi, P Luo, CJ Zhang, AF Molisch (2020). IEEE Transactions on Wireless Communications 19 (4), 2728-2741

[2]. Romano Fantacci, Benedetta Picano (2020): Federated learning framework for mobile edge computing networks. CAAI Trans. Intell. Technol. 5(1): 15-21

[3]. Heyam H. Al-Baity, Norah I. Al Showiman (2019). Towards Effective Service Discovery using Feature Selection and Supervised Learning Algorithms, *International Journal of Advanced Computer Science and Applications (JACSA)* Vol. 10 No. 5 pp. 191-200

[4]. A. Laachemi and D. Boughaci (2016), "A stochastic local search combined with support vector machine for Web services classification in International Conference on Advanced Aspects of Software Engineering (ICAASE), Constantine, pp. 9–16.

[5]. J. Liu, Z. Tian, P. Liu, J. Jiang, and Z. Li (2016) "An Approach of Semantic Web Service Classification Based on Naive Bayes," in IEEE International Conference on Services Computing (SCC), San Francisco, CA, pp. 356–362.

[6]. A. Syed Mustafa and Y. Kumara Swamy (2015), "Web Service classification using Multi-Layer Perceptron optimized with Tabu search," in IEEE International Advance Computing Conference (IACC), Banglore, pp. 290–294.

[7]. K. Bonawitz et al., (2019) "Towards federated learning at scale: System design, *in System Machine Language (SysML)*, 2019, to appear. [Online]. Available: https://arxiv.org/abs/1902.01046

[8]. H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas (2017) Communication-efficient learning of deep networks from decentralized data, *in Proc. Int. Conf. on AI and Statist*., Apr., 2017, pp. 1273–1282.

[9]. Jakub Konechy, H. Brendan McMahan, Daniel Ramage (2015), Federated Optimization: Distributed optimization beyond Datacenter *Google work*

[10]. Kunal Chandiramani, Dhruv Garg Maheswari N (2019), Performance Analysis of Distributed and Federated Learning Models on Private Data, *International Conference on Recent Trend International*

Conference On Recent Trends In Advanced Computing , ICRTAC Procedia Computer Science 165 (2019) 349–355

[11]. Q. Yang, Y. Liu, T. Chen, Y. Tong (2019), Federated machine learning: Concept and applications, *ACM Transactions on Intelligent Systems and Technology* (TIST) 10 (2) 12.

[12]. C. Song, T. Ristenpart, V. Shmatikov (2017), Machine learning models that remember too much, in: Proceedings of the *ACM SIGSAC Conference on Computer and Communications Security*, ACM, 2017, pp. 587-601. Conference (CEEC) pp. 57–60.

[13]. R. Shokri, M. Stronati, C. Song, V. Shmatikov (2017), Membership inference attacks against machine learning models, in: Security and Privacy (SP), *IEEE Symposium on*, *IEEE*, pp. 3{18.doi:https://doi.org/10.1109/SP.2017.41.

[14]. J. Konecny, H. B. McMahan, D. Ramage, and P. Richtarik (2016), Federated optimization: Distributed machine learning for on-device intelligence, arXiv preprint arXiv:1610.02527, 2016.

[15]. Kitchenham, B.A., Charters, S. (2009): Guidelines for performing systematic literature reviews in software engineering. version 2.3. EBSE Technical Report EBSE-2007- 01, Keele University, Keele, Staffordshire, United Kingdom

[16]. Yamada, M., Oda, T., Matsuo, K., & Barolli, L. (2016). Design of an IoT-Based E-learning Testbed. In 2016 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA) pp. 720–724.

[17]. Oda, T., Liu, Y., Yamada M., Matsuo, K., Ikeda, M., & Barolli, L. (2016). Performance Evaluation of an IoT-based e-Learning Testbed Considering OLSR Protocol in a NLoS Environment. In 2016 19th International Conference on Network-Based Information Systems (NBiS) pp. 451–457.

[18]. Mahmoud, Q. H., & Qendri, D. (2016). The Sensorian IoT platform. In 2016 13th IEEE Annual Consumer Communications Networking Conference (CCNC) pp. 286–287.

[19]. Raikar, M. M., Desai, P., & Naragund, J. G. (2016). Active Learning Explored in Open Elective Course: Internet of Things (IoT). In 2016 IEEE Eighth International Conference on Technology for Education (T4E) pp. 15–18.

[20]. Maksimović, M., Vujović, V., & Perišić, B. (2015). A custom Internet of Things healthcare system. In 2015 10th Iberian Conference on Information Systems and Technologies (CISTI) pp. 1–6.

[21]. Hentschel, K., Jacob, D., Singer, J., & Chalmers, M. (2016). Super sensors: Raspberry Pi Devices for Smart Campus Infrastructure.