



HAL
open science

Survey and enhancements on deploying LSTM recurrent neural networks on embedded systems

Ghalid Abib, Florian Castel, Nissrine Satouri, Hossam Afifi, Adel Mounir Said

► **To cite this version:**

Ghalid Abib, Florian Castel, Nissrine Satouri, Hossam Afifi, Adel Mounir Said. Survey and enhancements on deploying LSTM recurrent neural networks on embedded systems. ICC 2023: IEEE International Conference on Communications, IEEE, May 2023, Rome, Italy. pp.2034, 10.1109/ICC45041.2023.10278766 . hal-04173980

HAL Id: hal-04173980

<https://hal.science/hal-04173980v1>

Submitted on 1 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Can IIoT really exist? A Survey and a Practical Use Case Study

Florian Castel[†], Nissrine Satouri^{*†‡}, Ghalid Abib, and Hossam Afifi ,Adel Sareh Said[‡]
[†] Telecom SudParis,France
[‡] NTI,Egypt

Abstract—A real implementation of recurrent neural networks in low complexity IoT micro controllers is evaluated. The long short Term Memory algorithm is then adapted to fit into 32 bits devices and the back propagation algorithm is integrated inside. So we prove that IIoT can not only apply neural networks but also learn by themselves. Furthermore, a distributed LSTM architecture among different micro controllers is implemented and evaluated in feed forward conditions. The resulting Intelligent IoT architecture is proven to work efficiently and is compared to a classical gateway based learning platform. Results confirm that IIoT devices that learn by themselves are considered to be a real efficient solution that adds resilience to the architecture, reduces the storage and calculation costs and that is more energy friendly

Index Terms—Intelligent Iot, Machine learning, reinforcement learning techniques, LSTM.

I. INTRODUCTION

The development of Internet of Things (IoT) technologies has opened the gateway to a world smarter with a myriad of applications. Artificial intelligence (AI) has flourished due to the advent of the IoT [Pervasive AI for IoT applications: A Survey on Resource-efficient Distributed Artificial Intelligence]. AI is becoming more pervasive in the world of low-power embedded systems than ever before. One of the most recent applications that the world is interested in is reducing energy consumption and switching to green energy sources [A Smart Meter Infrastructure for Smart Grid IoT Applications].

On the one hand, and towards the smart energy grid, machine learning (ML) family technologies can be used to predict the energy consumption of customers/buildings/regions/cities. Thus, better smart power grid capacity planning can be achieved. On the other hand, there are many designs of digital energy meters that transmit their reading to a distributed system to run a Neural Network (NN) algorithm to loop on the prediction process.

In this work, we want to go further and test other types of more complex approaches to integrate new prediction algorithms on microcontroller (MC) with limited resources. Thus, our aim is to set up a NN on a MC: Espressif ESP32. Long Short-Term Memory (LSTM) has been chosen which is well adapted to the prediction of time series. In our case, it is not only a question of predicting the consumption of a smart electricity network. The goal of the work is building a simple edge processing platform to reduce the power consumption of prediction by adopting an MC running LSTM instead of

using a distributed server based system. The contributions of the paper are as follows: - Design and deploy a version of Backpropagation (BP) which is the neural network training distributed on MC. - Analysis of the quantification of weights on the quality of predictions and classifications.

The rest of the paper is organized as follows:.....

II. RELATED WORK

In this section we present the relevant studies in the area of UAV assisted new network generations.

III. RE VISITING THE LSTM ALGORITHM

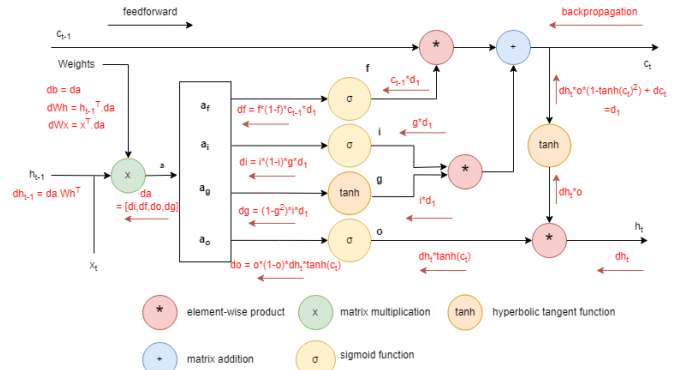


Fig. 1: Feedforward and backpropagation diagram in one LSTM cell

In [1] Feedforward and backpropagation are illustrated in Fig. 1. Expressions in red are the local gradients that are accumulated and propagated to previous time steps.

IV. REINFORCEMENT LEARNING APPROACH

Deep LSTM: Our choice goes to a hybrid LSTM network with 6 cells and 100 hidden layers, completed by a flat dense layer at the output. Inputs for the NN are anomalies and time taken in the ILP over the training set (2/3 of the 2 months of CDRs). The training labels correspond to the chosen action.

Results: T

A. Discussion

W

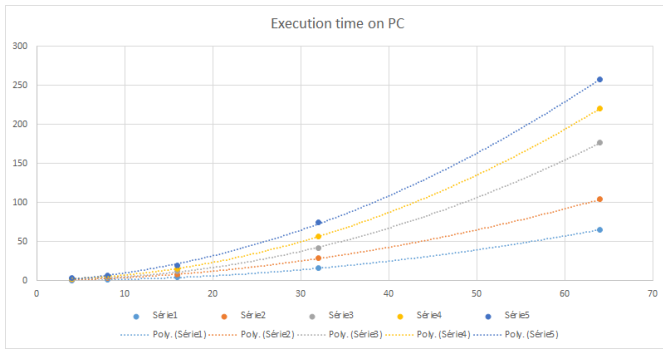


Fig. 2: Simple representation of case II

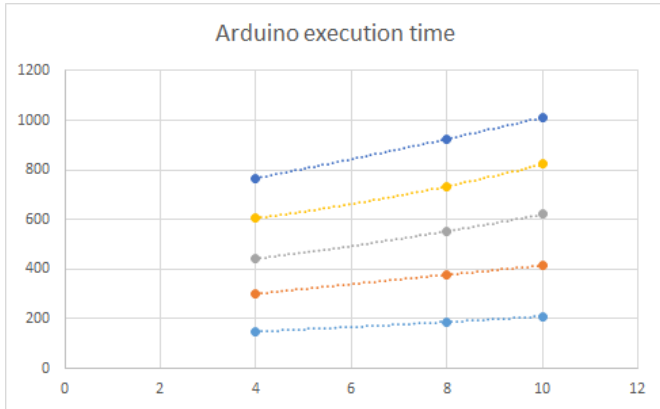


Fig. 3: Simple representation of case II

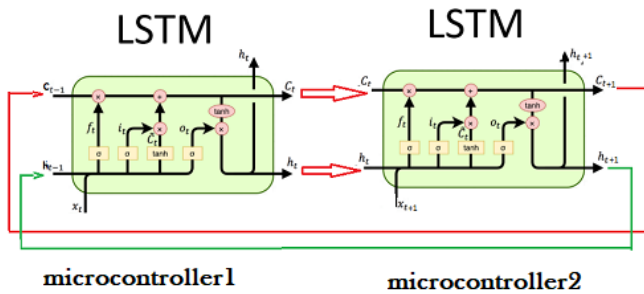


Fig. 4: Simple representation of case II

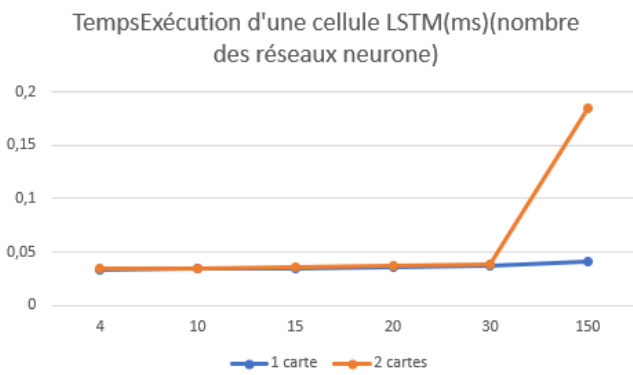


Fig. 6: execution time comparison

V. CONCLUSION AND FUTURE WORK

This paper presents a first approach to learn optimization ILP

A second deep learning based RL architecture is used and compared to the first one. It is based on deep LSTM networks. It learns faster than Q learning and reaches 80% of good actions.

Our fu

REFERENCES

- [1] P. Trakas, F. Adelantado, N. Zorba, and C. Verikoukis, "A quality of experience-aware association algorithm for 5g heterogeneous networks," 05 2017, pp. 1–6.