Dynamically reconfigurable data pipeline in the edge network

M.G.I.M. Nuwanthilaka (219376N) Supervised by Dr.Gayashan Amarasinghe

Introduction

Data pipelines are playing a significant role in emerging data science landscape as modern organizations ingest streams and batches of data with high variety, velocity and volume, from different sources. These pipelines mainly focus on extraction, data preparation, cleaning , transforming, training models and visualization. There are many data pipeline requirements and challenges in the field such as data cleansing on large scale data, security implementation to pipelines, visualization on large data sets etc. Apart from that, data pipelines need to comply with the security standards. As a result, the architectural support to implement these requirements is a major requirement.

Research problem

Nowadays, edge devices are resource rich compared to a few years back. There are tools/libraries developed in the industry which enable effective use of machine learning (ML) (ex: tensorflow mobile sdk) in the edge devices. With the introduction of such tools, the developers and the community have an opportunity execute ML tasks on the edge devices.

Most data pipelines are static with fixed configurations for all the models deployed on the pipeline. As a result, there exists an opportunity to optimise the configurations based on the constraints and requirements of the application. In this research, we focus on dynamically reconfiguring the architecture of data pipelines to optimally utilise the resources available in the edge and cloud devices.

The research work of Akoglu and Vargas-Solar [1],Raj et al. [4],Helu et al. [2],Konstantinou and Paton [3] and Renart et al. [5] have proposed edge frameworks. However, our research proposes a method investigating the possibilities of re-configurable architectures on the edge for data pipelines.

Research objectives

In this research, following three main objectives are addressed.

1. Dynamic reconfiguration of the pipeline architecture.
2. Optimal/near-optimal utilisation of edge resources to support proposed architecture.
3. Evaluation of algorithms that are benefited from this approach (performance evaluation).

Proposed method

First, resources of the edge devices and the requirements of the applications will be analysed to identify the key parameters which should be given the priority and then an architecture which could be able to dynamically reconfigure based on the resources will be proposed. The proposed architecture will be implemented on Raspberry Pi and/or Android mobile devices. Then those configurations will be optimized to support maximum utilization of the resources.

In the evaluation phase, a selected set of ML algorithms will be analysed to address how the proposed architecture will benefit for these algorithms to perform better. Apart from the proposed reconfigurable architecture, another outcome of the proposed research is to produce a classification of ML algorithms that can benefit from the proposed architecture based on their characteristics.

References

[1] Ali Akoglu and Genoveva Vargas-Solar. Putting Data Science Pipelines on the Edge. pages 1–13, 2021. URL `http://arxiv.org/abs/2103.07978`.

[2] Moneer Helu, Timothy Sprock, Daniel Hartenstine, Rishabh Venketesh, and William Sobel. Scalable data pipeline architecture to support the industrial internet of things. *CIRP Annals*, 69(1): 385–388, 2020. ISSN 17260604. doi: 10.1016/j.cirp.2020.04.006.

[3] Nikolaos Konstantinou and Norman W. Paton. Feedback driven improvement of data preparation pipelines. *Information Systems*, 92, 2020. ISSN 03064379. doi: 10.1016/j.is.2019.101480.

[4] Aiswarya Raj, Jan Bosch, Helena Holmstrom Olsson, and Tian J. Wang. Modelling Data Pipelines. *Proceedings - 46th Euromicro Conference on Software Engineering and Advanced Applications, SEAA 2020*, pages 13–20, 2020. doi: 10.1109/SEAA51224.2020.00014.

[5] Eduard Gibert Renart, Daniel Balouek-Thomert, and Manish Parashar. An edge-based framework for enabling data-driven pipelines for IoT systems. *Proceedings - 2019 IEEE 33rd International Parallel and Distributed Processing Symposium Workshops, IPDPSW 2019*, pages 885–894, 2019. doi: 10.1109/IPDPSW.2019.00146.