Data Markets: A Multi-Agent System Platform for Cooperative and Competitive Distributed Machine Learning

Hajar Baghcheband[1] , Carlos Soares[1] and Luís Paulo Reis[1]

[1] Faculdade de Engenharia, Universidade do Porto (FEUP),
{h.baghcheband,csoares,lpreis}@fe.up.pt

Abstract. Fast-growing of the generated data and costly analysis of huge data attract researchers to investigate distributed systems and distributed learning. Knowledge extraction and efficiently utilize distributed data in such complex and noisy system become more challengeable field. Multi-agent systems (MAS) offer an architecture for distributed problem solving. Distributed data mining (DDM) and machine learning (DML) algorithms focus on one class of such distributed problem solving tasks—analysis and modeling of distributed data. Whereas data analysis is costly task, data can be considered as a product and from perspective of data marketing, economic strategies are noticeable issue. In this research, multi-agent system platform for cooperative and competitive distributed machine learning will be designed to reach best data trading in distributed data market. It is expected that the proposed data trading model and market strategy can be applied to different domains of marketing such as smart cities, smart transportation.

Keywords: Multi-Agent System, Distributed Data Mining and Machine Learning, Data Market, Data Trading Model.

1 Introduction

The produced data explode due to the emerging technologies and applications that pervade everywhere in our daily lives, including internet of things applications [1] and distributed data source makes it harder. With this exponential growth, how to efficiently utilize, analyse and extract the data become critical issues and the development of a big data market that enables suitable data trading is needed. By considering data as a kind of commodity in a digital market, the data owners and consumers are able to connect with each other, sharing and further increasing the utility of data [2], [3]. Whereas data is a vital item, "new money", with special characteristics such as variability, velocity, and complexity, traditional pricing model must be improved [2].

Regard to the shortage of feasible protocols, existing data trading markets are still in the initial stages. In this market, several challenges need to be addressed, such as determining proper pricing for the data trading, designing an adequate platform and schemes to enable the maximization of the utility of both-side participants with efficiency and

privacy preservation, and protecting the traded data from being resold (data copyright protection) to maintain the value of the data [2].

2 Motivation, Goals and Hypothesis

To analyzing data which is essential to monitor critical smart-world infrastructures [4], machine learning, an approach for extracting accurate information, meets two challenges, huge number of data and centralized storing source. Multi-agent systems (MASs) which deal with complex applications is one distributed nature [5]. Different distributed approaches fall into the area of distributed systems, where a number of entities work together to cooperatively or competitively solve problems[6]. Combining MAS with distributed data mining (DDM) for data intensive applications is appealing to deal with challenges in analyzing distributed data, offer many learning algorithmic solutions with best performance among several work stations and perform different data mining operations in a fundamentally distributed manner [5]. In many applications the most important thing is to inform its decision making [7].

Applying learning and data mining are cost task for service and data providers, in pricing market, the basic goal of both static or dynamic pricing is to incentivize users to adjust their demands so as to maximize the overall resource utilization. In this situation, data becomes a product and distributed systems with capabilities of distributed data mining and decision support as problem solving systems become markets.

The goals of this research are to investigate the development and the empirical analysis of a distributed multi-agent systems as platform, model a market of machine learning agents that trade data and improve the quality of the models they develop with reasonable costs and time. It is expected that by combination of these knowledge, the proposed data trading modeling can be designed to optimize the performance of learning application in distributed environments.

References

1. H. Li, K. Ota, and M. Dong, “Learning IoT in Edge: Deep Learning for the Internet of Things with Edge Computing,” IEEE Netw., vol. 32, no. 1, pp. 96–101, 2018.
2. F. Liang, W. Yu, D. An, Q. Yang, X. Fu, and W. Zhao, “A Survey on Big Data Market: Pricing, Trading and Protection,” IEEE Access, vol. 6, pp. 15132–15154, 2018.
3. Y. Zhao, Y. Yu, Y. Li, G. Han, and X. Du, “Machine learning based privacy-preserving fair data trading in big data market,” Inf. Sci. (Ny)., vol. 478, pp. 449–460, 2019.
4. W. Yu et al., “A Survey on the Edge Computing for the Internet of Things,” IEEE Access, vol. 6, pp. 6900–6919, 2017.
5. J. C. Da Silva, C. Giannella, R. Bhargava, H. Kargupta, and M. Klusch, “Distributed data mining and agents,” Eng. Appl. Artif. Intell., vol. 18, no. 7, pp. 791–807, 2005.
6. L. PANAIT and S. LUKE, “Cooperative Multi-Agent Learning: The State of the Art,” Auton. Agent. Multi. Agent. Syst., vol. 11, no. 3, pp. 387–434, 2005.
7. S. V. Albrecht and P. Stone, “Autonomous agents modelling other agents: A comprehensive survey and open problems,” Artif. Intell., vol. 258, no. January, pp. 66–95, 2018.