CNP-MLDM: Contract Net Protocol for Negotiation in Machine Learning Data Market

Hajar Baghcheband[1,2,*,†], Carlos Soares[1,2,3,†] and Luis Paulo Reis[1,2,†]

[1]FEUP - Faculty of Engineering University of Porto, Porto, Portugal
[2] LIACC-Artificial Intelligence & Computer Science Laboratory(member of LASI LA), Porto, Portugal
[3]Fraunhofer AICOS Portugal, Porto, Portugal

Abstract

The Machine Learning Data Market (MLDM), which relies on multi-agent systems, necessitates robust negotiation strategies to ensure efficient and fair transactions. The Contract Net Protocol (CNP), a well-established negotiation strategy within Multi-Agent Systems (MAS), offers a promising solution. This paper explores the integration of CNP into MLDM, proposing the CNP-MLDM model to facilitate data exchanges. Characterized by its task announcement and bidding process, CNP enhances negotiation efficiency in MLDM. This paper describes CNP tailored for MLDM, detailing the proposed protocol following experimental results.

Keywords

Machine Learning, Data Market, Negotiation, Contract Net Protocol

1. Introduction

The Machine Learning Data Market (MLDM), a data market framework based on multi-agent systems [1, 2, 3, 4], is designed to facilitate data exchange among agents to enhance their predictive performance. Each agent initially builds a model using local training data and assigns a value to newly collected data for potential trading. Agents then engage in negotiations to buy or sell data, setting prices based on the data valuation methods and their respective budgets. Through these transactions, agents aim to improve their learning models and overall performance by incorporating exchanged data. The framework also includes mechanisms for performance evaluation after exchanging data to assess the impact on predictive accuracy and budget.

While the original negotiation strategy facilitated data exchange, it lacked efficiency. To address this, we introduced a tailored version of the Contract Net Protocol (CNP) [5], a well-established MAS negotiation protocol [6]. This new approach, termed CNP-MLDM, aims to optimize data exchange processes specifically within the MLDM framework.

2. CNP-MLDM

In the MLDM, agents engage in dynamic data exchanges to enhance predictive modeling performance. Seller Agents (SAs) and Buyer Agents (BAs) participate in a structured negotiation facilitated by the CNP-MLDM protocol. Each agent develops models based on local datasets, selects traded sets, and sets prices accordingly. Through CNP-MLDM, agents announce their intentions to sell data daily (iteration), including the price and data valuation of the traded set(see Figure 1a). Buyers analyze received and unexpired offers, calculating the worth according to price and data valuation (DV) $worth = (DV/price) * willingToBuy$. Buyers then rank offers based on their worth and select the best one. If the buyer's budget is insufficient, it suggests a new price based on *willingToBuy and data value*. After

Discovery Science - Late Breaking Contributions 2024
*Corresponding author.
†These authors contributed equally.
h.baghcheband@fe.up.pt (H. Baghcheband); csoares@fe.up.pt (C. Soares); lpreis@fe.up.pt (L. P. Reis)
0000-0003-1587-4635 (H. Baghcheband); 0000-0003-4549-8917 (C. Soares); 0000-0002-4709-1718 (L. P. Reis)
© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
CEUR-WS.org/Vol-3928/paper_173.pdf

To: Broadcast
From: Sender`s Name
Type: Sell data
Description: value of traded data
Price
Expiration date: day (Iteration)

(a) Task Announcement Message

To: Sender`s Name
From: Buyer`s Name
Type: Buy data
Description: Accepted/ Suggested Price
Ask for additional Info. (if any)
Expiration date: day (Iteration)

(b) Bid Announcement Message

choosing an offer, the buyer agent sends a response to the seller with the accepted or suggested price and any other required information (see Figure 1b). The seller agent assesses received messages from buyers. If a new price is suggested, the seller decides on the reduced traded data (size/quality). The seller then sends the traded data to the buyer, who appends the new set to its training data and evaluates model performance to measure the effect of the data exchange. If the buyer's performance improves, it can add the seller to a trust list for future transactions.

The performance of the MLDM framework using the Gain-Shapley Data Value (GDSV) method for data valuation was examined on the OpenMl-CC18 [7] (a set of classification datasets) with the K-Nearest Neighbors (KNN) algorithm. The results compare three scenarios: GDSV, No Exchange, and Single Agent. The GDSV scenario significantly outperforms the No Exchange scenario, with the performance gap widening as data collection increases. The Single Agent scenario, representing the performance with complete data access, consistently shows the highest performance. However, the GDSV scenario approaches this maximum boundary, indicating that strategic data exchanges enable agents to approximate the performance of a single agent collectively. Overall, the results underscore the benefits of the MLDM framework with GDSV and Contract Net Protocol (CNP) negotiation, demonstrating that intelligent data trading can significantly improve predictive performance(see Figure 2).

3. conclusion

This framework enhances the negotiation strategy in MLDM, fostering a cooperative marketplace where agents collectively improve their learning models for better overall performance. Future work will involve adjusting the scenario to the real world based on complex scenarios like updated prices from buyers and the bidding mechanism for sellers, taking into account changes in the size or quality of the traded data. Additionally, a trust list can be implemented to manage competitive relationships among agents.

Acknowledgments

This work was financially supported (or partially financially supported) by Base Funding – UIDB/00027/2020 of the Artificial Intelligence and Computer Science Laboratory – LIACC - funded by national funds through the FCT/MCTES (PIDDAC) and by a PhD grant from Fundação para a Ciência e Tecnologia (FCT), reference SFRH/BD /06064/2021 and Agenda "Center for Responsible AI'', nr. C645008882-00000055, investment project nr. 62, financed by the Recovery and Resilience Plan (PRR) and by European Union - NextGeneration EU. The computational resources of Google Cloud Platform were provided by the project CPCA-IAC/AF/594904/2023.

References

[1] H. Baghcheband, C. Soares, L. P. Reis, Machine learning data markets: Evaluating the impact of data exchange on the agent learning performance, 2023. Paper presented at the 22nd Portuguese Conference on Artificial Intelligence, 5–8 September 2023.

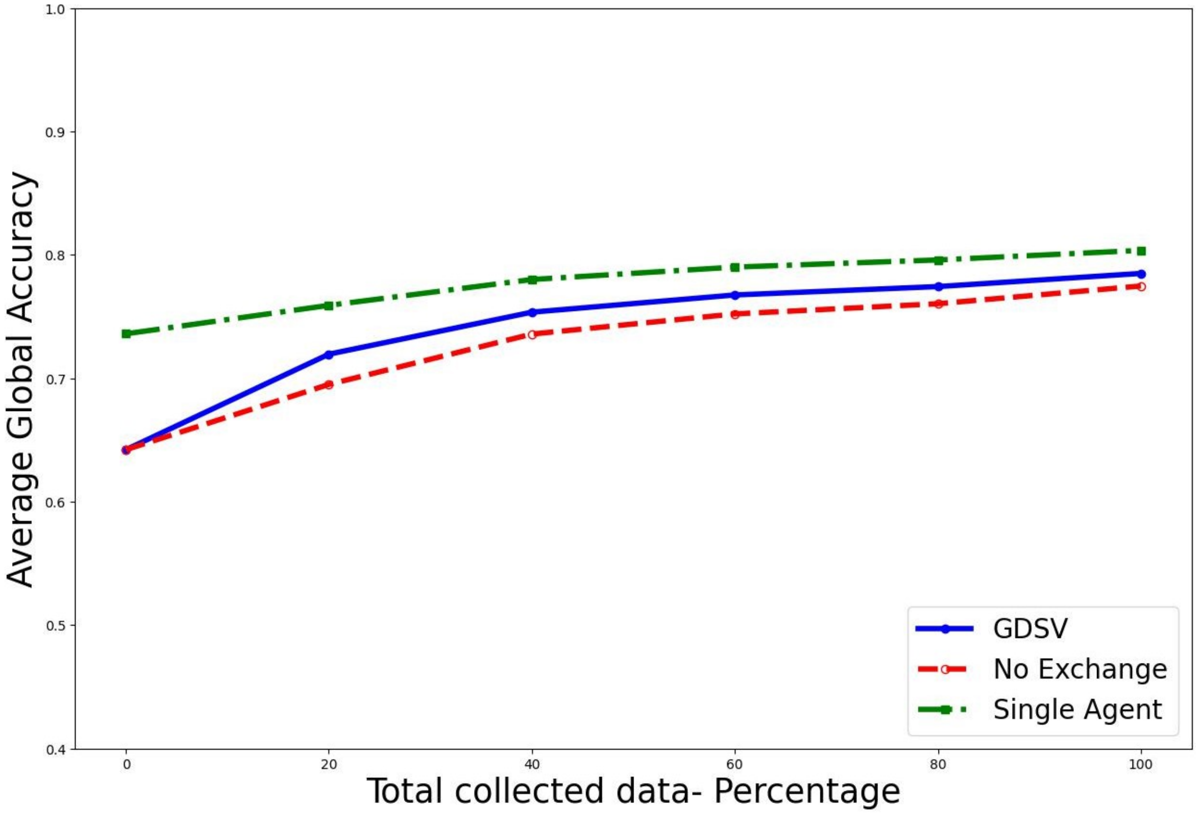

Figure 2: Experimental Results according to different scenarios (Single Agent, MLDM, No Exchange), GDSV as data valuation and CNP as negotiation strategy

[2] H. Baghcheband, C. Soares, L. P. Reis, Machine learning data markets: Trading data using a multi-agent system, in: 2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), IEEE Computer Society, Los Alamitos, CA, USA, 2022, pp. 450–457.

[3] H. Baghcheband, C. Soares, L. Reis, Mldm: Machine learning data market based on multi-agent systems, IEEE Internet Computing (2024) 1–7.

[4] H. Baghcheband, C. Soares, L. P. Reis, Shapley-based data valuation method for the machine learning data markets (mldm), in: Foundations of Intelligent Systems, Springer Nature Switzerland, Cham, 2024, pp. 170–177.

[5] R. G. Smith, The contract net protocol: High-level communication and control in a distributed problem solver, IEEE Transactions on Computers C-30 (1981) 372.

[6] D. Sun, J. Wu, Multi-agent coordination based on contract net protocol, 2009 International Symposium on Intelligent Ubiquitous Computing and Education, IUCE 2009 (2009) 353–357.

[7] B. Bischl, G. Casalicchio, M. Feurer, F. Hutter, M. Lang, R. G. Mantovani, J. N. van Rijn, J. Vanschoren, Openml benchmarking suites, arXiv:1708.03731v2 [stat.ML] (2019).