Toward Approximate Intelligence

Approximate Query Engines & Approximate Data Exploration

Dominik Ślęzak
Faculty of Mathematics, Informatics and Mechanics
University of Warsaw
ul. Banacha 2, Warsaw, Poland
slezak@mimuw.edu.pl

ABSTRACT

The AI methods are regaining a lot of attention in the areas of data analytics and decision support. Given the increasing amount of information and computational resources available, it is now possible for intelligent algorithms to learn from the data and assist humans more efficiently. Still, there is a question about the goals of learning and a form of the resulting data-driven knowledge.

It is evident that humans do not operate with precise information in decision-making and, thus, it might be unnecessary to provide them with complete outcomes of analytical processes. Consequently, the next question arises whether approximate results of computations or results derived from the approximate data could be delivered more efficiently than their standard counterparts.

Such questions are analogous to the ones about precision of calculations conducted by machine learning and KDD methods, whereby heuristic algorithms could be boosted by letting them rely on approximate computations. This leads us toward discussion of the importance of approximations in the areas of machine intelligence and business intelligence and, more broadly, the meaning of approximate derivations for various aspects of AI. In this talk, this discussion is supported by four industry-related case studies¹:

- (1) The approximate database engine based on the paradigms of rough-granular computing applied in the area of cyber-security analytics [6], [3], [5];
- (2) The similarity-based feature engineering methodology embedded into an HR support system working with heterogeneous information sources [4];
- (3) The ensemble-based attribute approximation approach that will be used in the area of online health support [2];
- (4) The process of approximate data generation that will be used for tuning an online gaming coaching platform [1].

¹The first case study refers entirely to the author's work for Security On-Demand (<https://www.securityondemand.com/>). The work on the second case study was co-financed by the EU Smart Growth Operational Programme 2014-2020 under the Innovation Voucher project POIR.02.03.02-14-0009/15-00. The work on the third case study is co-financed by the EU Smart Growth Operational Programme 2014-2020 under the project POIR.01.01.01-00-0831/17-00. The work on the fourth case study is co-financed by the EU Smart Growth Operational Programme 2014-2020 under the GameINN project POIR.01.02.00-00-0184/17-00.

This paper is published under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW'18 Companion, April 23–27, 2018, Lyon, France

© 2018 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC BY 4.0 License.

ACM ISBN 978-1-4503-5640-4/18/04.
<https://doi.org/10.1145/3184558.3193133>

CCS CONCEPTS

• **Information systems** → **Database management system engines**; *Decision support systems*; • **Computing methodologies** → *Artificial intelligence*; **Machine learning approaches**; • **Security and privacy** → *Intrusion/anomaly detection and malware mitigation*; • **Applied computing** → *Health informatics*;

KEYWORDS

Approximate Data Analytics; Approximate Database Engines; Approximate Data Exploration; Approximate Feature Selection

ACM Reference Format:

Dominik Ślęzak. 2018. Toward Approximate Intelligence: Approximate Query Engines & Approximate Data Exploration. In *The 2018 Web Conference Companion, April 23–27, 2018, Lyon, France*. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/3184558.3193133>

REFERENCES

- [1] 2018. eSensei. (2018). <http://esensei.net/>
- [2] 2018. Ovufriend. (2018). <https://ovufriend.pl/>
- [3] Agnieszka Chądzyńska-Krasowska, Paweł Betliński, and Dominik Ślęzak. 2017. Scalable Machine Learning with Granulated Data Summaries: A Case of Feature Selection. In *Proceedings of ISMIS 2017* (Warsaw). 519–529.
- [4] A. Janusz, S. Stawicki, M. Drewniak, K. Ciebiera, D. Ślęzak, and K. Stencel. 2018. How to Match Jobs and Candidates – A Recruitment Support System Based on Feature Engineering and Advanced Analytics. In *Proceedings of IPMU 2018* (Cadiz).
- [5] D. Ślęzak, A. Chądzyńska-Krasowska, J. Holland, P. Synak, R. Glick, and M. Perkowski. 2017. Scalable Cyber-security Analytics with a New Summary-based Approximate Query Engine. In *Proceedings of BigData 2017* (Boston). 1840–1849.
- [6] Dominik Ślęzak, Rick Glick, Paweł Betliński, and Piotr Synak. 2018. A New Approximate Query Engine Based on Intelligent Capture and Fast Transformations of Granulated Data Summaries. *Journal of Intelligent Information Systems* 50, 2 (2018), 385–414.