Discovering the misconceptions that influence the learning of Machine Learning

Author

Iulia Angelescu

Supervisors

Gosia Migut
Marcus Specht

Affiliations

This research is conducted as part of
TU Delft's course entitled "Research Project"

A research that analyses the existing literature on misconceptions in the context of Machine Learning and directly applies the findings to the course of Machine Learning offered at the Technical University of Delft as part of the Computer Science and Engineering Bachelor's curriculum.

1. Introduction

The Machine Learning field presents high demands but lacks the engineering expertise required to meet the needs of the domain.[1]
According to already existing research, between 200,000 and 300,000 engineers of the 18 million software developers in the world possess the qualifications to engage with AI and ML methods. [2]

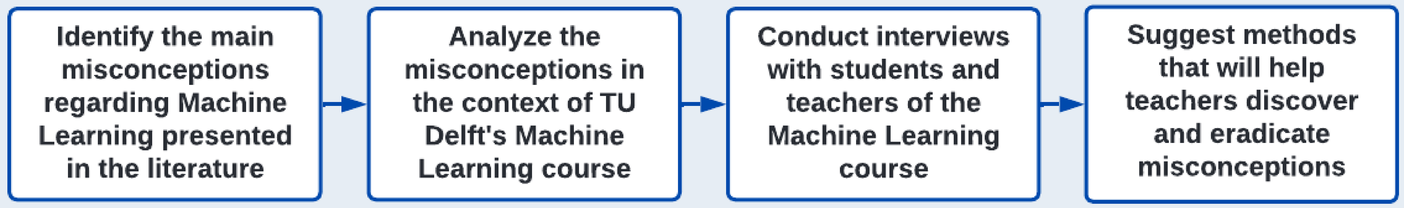
2. Objectives

Which are the misconceptions that influence learning of Machine Learning?

The research unveils:

- the main misconceptions regarding Machine Learning
- the origin of these misconceptions
- how these misconceptions can be eradicated
- how can teachers detect other misconceptions

3. Methodology



4. Results

Literature results

Misconceptions fall under two categories

1. Misconceptions with respect to the importance and understanding of the field of Machine Learning [3]
2. Misconceptions regarding specific models of Machine Learning and their applicability.[3]

The first group of misconceptions includes:

- ML algorithms can be applied without having expertise in the field
- ML represents computers' ability to improve their performance without human help
- ML and AI are considered to be two distinct technologies that attempt to accomplish different objectives

Interviews' results

18 former students of the course of ML were interviewed and misconceptions regarding the following topics were pointed out:

- **61%** Principal Component Analysis
- **50%** Cross-Validation
- **100%** Gradient Descent and Stochastic Gradient Descent

With the help of 3 teachers of the ML course, different methods that could help eradicate the above misconceptions were created:

- an exercise that points out the differences between Principal Component Analysis and Feature Selection
- an exercise that highlights the values of using Cross-Validation

5. Conclusion

- Misconceptions are primarily caused by an imbalance between the practical activities and the theoretical information that is covered during a course [2].
- The proposed approaches can be used to eradicate the mentioned misconceptions and to discover new ones.

6. Limitations & further work

- The number of interviewed students can be increased in order to obtain a more accurate generalization.
- The study can continue by analyzing the student's former ML assignments.

RELATED LITERATURE

[1] A. J. Ko, We need to learn how to teach machine learning, medium.com/bits-and-behavior, Ed., https://medium.com/bits-and-behavior/we-need-to-learn-how-to-teach-machine-learning-acc78bac3ff8, Aug. 2017
[2] H. Heuer, J. Jarke, and A. Breiter, "Machine learning in tutorials – Universal applicability, under-informed application, and other misconceptions", Big Data & Society, vol. 8, no. 1, p. 20 539 517 211 017 593,2021. doi: 10 . 1177 / 20539517211017593. [Online]. Available: https : / / doi . org / 10 . 1177 /20539517211017593
[3] E. Marx, T. Leonhardt, D. Baberowski, and N. Bergner, "Using matchboxes to teach the basics of machine learning: An analysis of (possible) misconceptions", Proceedings of Machine Learning Research, vol. 170, K. M. Kinnaird, P. Steinbach, and O. Guhr, Eds., pp. 25–29, Aug. 2022. [Online]. Available: https://proceedings.mlr.press/v170/marx22a.html

TU Delft