By: Elmedin Burnik
Supervisor(s): dr. Gosia Migut, prof. dr. Marcus Specht

The order of competences that students need to learn in ML

1 Background (figure 1, 2)

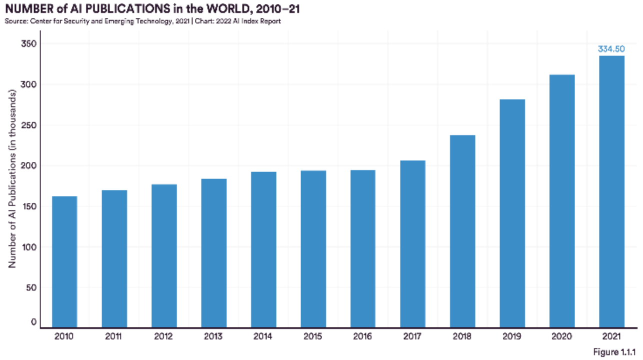

Figure 1: Number of AI publications in the world between 2010 and 2021 from the centre of Security and Technology AI Index Report 2022 [1].

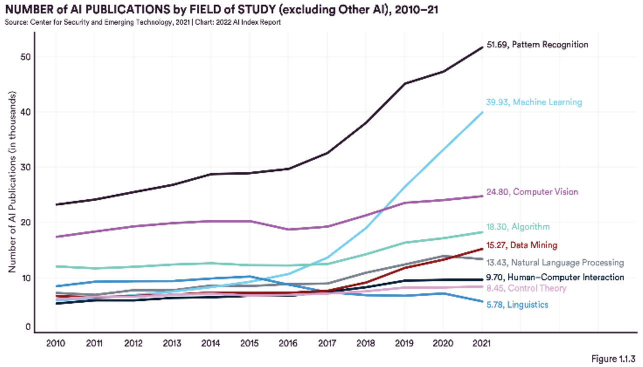

Figure 2: Number of AI publications by field of study between 2010 and 2021 from the centre of Security and Technology AI Index Report 2022 [1]. Showcasing the rise of ML publications within AI.

2 Methodology

The main research question:
"What is the order of competences that students need to learn in ML?"

1. Definition of <u>Competence</u>: *"A combination of attributes in terms of knowledge and its application, skills, responsibilities and attitudes."*
2. Gather competences from university course guides.
3. Perform questionnaire with academics ML teachers.
4. Use comparison methods to find resulting order of competences.

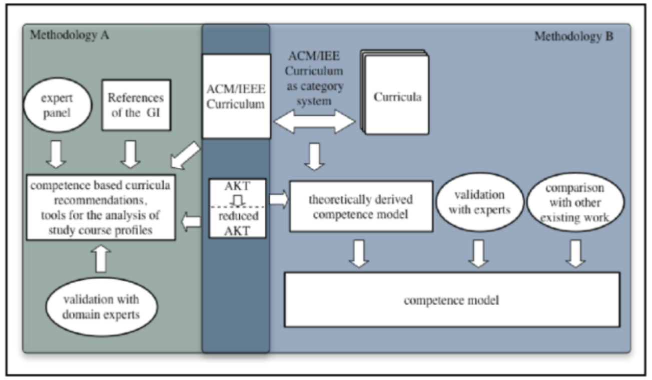

Figure 3. Methodologies for competence model development [2].

3 Analysis and Experiment on finding the order of Competences

- 2 methods used (fig. 3):
    1. <u>Literature Analysis</u>
    2. <u>Experimental Approach</u>

Kendall-τ distance method used throughout [3].

1. **Literature Analysis:**
- Course guides from different universities analyzed and compared to each other.
- Paper from Danyluk et al. results compared and combined with course guide competences [4].
- Paper by Von Wagenheim and another by Leidig that had established an order used for generating our own order [5-6].

2. **Experimental Approach**
- Questionnaire made, asked to order a randomized no. of competences.
- Answered by ML academics teachers.
- Results from the different teachers compared.

Results from 1 are compared and a final ordered list is established in table 1.

4 Results

1	**Characterize and differentiate between different classes of machine learning models (e.g., geometric, probabilistic) and tasks (e.g., classification, clustering, regression).**
11	**Provide an appropriate performance metric for evaluating ML algorithms/tools for a given problem.**
9	**Express formally the representational power of models learned by an algorithm and relate that to issues such as expressiveness and overfitting.**
2	**Derive a (current) learning algorithm from first principles and/or justify a (current) learning algorithm from a mathematical, statistical, or information-theoretic perspective.**
3	**Apply appropriate empirical evaluation methodology to compare ML algorithms/tools to each other.**
8	**Exhibit knowledge of methods to mitigate the effects of overfitting and curse of dimensionality in the context of ML algorithms.**
6	**Implement ML programs from their algorithmic specifications.**
13	**Explain the concept of and identify (implicit) bias in data and ML algorithms**
12	**Consider and evaluate the possible effects – both positive and negative – of decisions arising from ML conclusions.**
4	**Compare differences in interpretability of learned models.**
10	**Select and apply a broad range of ML tools/implementations to real data.**
5	**Apply appropriate empirical evaluation methodology to assess the performance of a ML algorithm/tool for a problem.**
7	**Understand and implement expertise in the Python programming language and its statistical and numerical libraries.**

Table 1. The resulting ordered list of competences. The numbered competences remain the same as in table 1 in the paper.

5 Future work

- Dissecting results into smaller sub-competences.
- Unique order of sub-competences within each competence.

[1] Daniel Zhang, Nestor Maslej, Erik Brynjolfsson, John Etchemendy, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Michael Sellitto, Ellie Sakhaee, Yoav Shoham, Jack Clark, and Raymond Perrault, "The AI Index 2022 Annual Report," AI Index Steering Committee, Stanford Institute for Human-Centered AI, Stanford University, March 2022. https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf
[2] Brinda, T., Reynolds, N., Romeike, R., & Schwill, A. (2015). KEYCIT 2014. Beltz Verlag.
[3] KENDALL, M. G. (1938). A NEW MEASURE OF RANK CORRELATION. Biometrika, 30(1–2), 81–93. https://doi.org/10.1093/biomet/30.1-2.81
[4] Danyluk, A., & Buck, S. (2019). Artificial Intelligence Competencies for Data Science Undergraduate Curricula. Proceedings of the AAAI Conference on Artificial Intelligence, 33, 9746–9747.
[5] Gresse Von Wangenheim, C., Marques, L. S., & Hauck, J. C. R. (2020). Machine Learning for All – Introducing Machine Learning in K-12. Machine Learning for All – Introducing Machine Learning in K-12. https://doi.org/10.31235/osf.io/wj5ne
[6] Leidig, P. M., & Cassel, L. (2020). ACM Taskforce Efforts on Computing Competencies for Undergraduate Data Science Curricula. Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education. https://doi.org/10.1145/3341525.3393962

TU Delft