Tensor Decomposition for Student Success Prediction Models in Hands-on Cybersecurity Exercises

Julia Scott, *Lewis & Clark College*
Jens Mache, *Lewis & Clark College*
Richard Weiss, *The Evergreen State College*
Contact: juliascott@lclark.edu

Cybersecurity is an ever-evolving field that demands more workers and a wider array of knowledge every year. As such, cybersecurity education remains essential — not just for professionals, but for developers and non-technical roles as well. Due to this, hands-on cybersecurity exercises, such as the ones in the eduRange platform, are increasingly important. EduRange aims to be a flexible, intuitive cybersecurity platform that allows instructors to tailor pre-existing scenarios to their classes' needs. However, when students become stuck or frustrated, learning grinds to a halt. To combat this discouragement, we want to create a semi-automated hint system that can consistently identify struggling students. Such a hint system, however, requires a large quantity of data, which can be difficult to obtain through classroom testing alone.

As such, we explored creating synthetic data. We used a sample dataset and stored attempt accuracy in a three dimensional tensor with dimensions students, questions, and attempts. We then used tensor decomposition to fill in gaps in the dataset, a process called densification. Our primary objective was to optimize the tensor decomposition to obtain the most accurate possible densification. The results showed that to obtain the greatest accuracy, we should use rank-1 tensors and fill in logical extra data points. Additionally, we found that generic tensor decomposition may not be sufficient for boolean data. In future work, we plan to use boolean tensor decomposition to improve our results.

Keywords: Cybersecurity Education; Intelligent Tutoring Systems; Synthetic Data; Tensor Decomposition

DOI: https://doi.org/10.1145/3641555.3705052