

Mining Subgroups with Exceptional Transition Behavior

(Abstract)

Florian Lemmerich^{1,5}, Martin Becker², Philipp Singer^{1,5},
Denis Helic³, Andreas Hotho^{2,4}, and Markus Strohmaier^{1,5}

¹ GESIS, Cologne, Germany

{florian.lemmerich, philipp.singer, markus.strohmaier}@gesis.org

² University of Würzburg, Germany

{becker, hotho}@informatik.uni-wuerzburg.de

³ Graz University of Technology, Graz, Austria

dhelic@tugraz.at

⁴ L3S Research Center, Hannover, Germany

⁵ University of Koblenz-Landau, Mainz, Germany

Abstract. In this talk, we present a recently developed method for detecting interpretable subgroups with exceptional transition behavior in sequential data [1]. Potential applications of this technique include, e.g., studying human mobility or analyzing the behavior of internet users. To approach this task, we extend exceptional model mining. Exceptional model mining provides a framework for mining interpretable data subsets with unusual interactions between a set of target attributes considering a user-chosen model class. However, previously investigated model classes cannot capture transition behavior. Thus, we introduce first-order Markov chains as a novel model class for exceptional model mining and present a new interestingness measure that quantifies the exceptionality of transition subgroups. The measure compares the distance between the Markov transition matrix of a subgroup and the respective matrix of the entire data with the distance of random dataset samples. In addition, our method can be adapted to find subgroups that match or contradict given transition hypotheses. We demonstrate that our method is consistently able to recover subgroups with exceptional transition models from synthetic data and illustrate its potential in two application examples. Our work is relevant for researchers and practitioners interested in detecting exceptional transition behavior in sequential data.

References

1. Lemmerich, F., Becker, M., Singer, P., Helic, D., Hotho, A., Strohmaier, M.: Mining subgroups with exceptional transition behavior. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2016), http://www.kdd.org/kdd2016/papers/files/Paper_185.pdf