

# Harnessing the Strengths of Anytime Algorithms for Constant Data Streams<sup>\*</sup>

Philipp Kranen and Thomas Seidl

Data Management and Exploration Department  
RWTH Aachen University, Germany  
{kranen, seidl}@cs.rwth-aachen.de

**Abstract.** Anytime algorithms have been proposed for many different applications e.g. in data mining. Their strengths are the ability to first provide a result after a very short initialization and second to improve their result with additional time. Therefore, anytime algorithms have so far been used when the available processing time varies, e.g. on varying data streams. In this paper we propose to employ anytime algorithms on constant data streams, i.e. for tasks with constant time allowance. We introduce two approaches that harness the strengths of anytime algorithms on constant data streams and thereby improve the overall quality of the result with respect to the corresponding budget algorithm. We derive formulas for the expected performance gain and demonstrate the effectiveness of our novel approaches using existing anytime algorithms on benchmark data sets.

The goal that was set and reached in this paper is to improve the quality of the result over that of traditional budget approaches, which are used in an abundance of stream mining applications. Using anytime classification as an example application we show for SVM, Bayes and nearest neighbor classifiers that both our novel approaches improve the classification accuracy for slow and fast data streams. The results confirm our general theoretic models and show the effectiveness of our approaches. The simple yet effective idea can be employed for any anytime algorithm along with a quality measure and motivates further research in e.g. classification confidence measures or anytime algorithms.

## Acknowledgments

This work has been supported by the UMIC Research Centre, RWTH Aachen University.

## Reference

1. Kranen, P., Seidl, T.: Harnessing the Strengths of Anytime Algorithms for Constant Data Streams. *Data Mining and Knowledge Discovery* (2009) DOI: 10.1007/s10618-009-0139-0

---

<sup>\*</sup> This is an extended abstract of an article published in the *Data Mining and Knowledge Discovery* journal [1].