University of Zurich[UZH]

Department of Informatics

University of Zürich
Department of Informatics
Binzmühlestr. 14
CH-8050 Zürich
Phone. +41 44 635 43 11
Fax +41 44 635 68 09
www.ifi.uzh.ch/dbtg

UZH, Dept. of Informatics, Binzmühlestr. 14, CH-8050 Zürich

Joana Cieri
joana.cieri@uzh.ch

Prof. Dr. Michael Böhlen
Professor
Phone +41 44 635 43 33
Fax +41 44 635 68 09
boehlen@ifi.uzh.ch

Zürich, 21. Dezember 2023

BSc Thesis
Topic: Implementing and Investigating Entropy Rate Estimation Algorithms of Finite Numeric Time Series

Determining the exact entropy rate of time series poses impractical challenges because of the unknown probability distribution of the underlying process and the exponential runtime complexity, even when these probabilities are known. For this reason, the entropy rate is estimated rather than precisely computed. Estimating the entropy rate of time series is a fundamental problem in information theory, as highlighted by Cover [1]. Actually, there is a whole range of work on entropy rate estimators and the have all been proven to converge to the true entropy rate for the length of the time series going to infinity. However, real-world time series are finite and for this scenario the estimators show different behaviors [3].

In this thesis work, the main goal is to implement and investigate entropy rate estimators based on block entropies [3] and Lempel-Ziv algorithms [1, 2, 3]. Specifically, the focus is on examining the effectiveness of these entropy rate estimators in both ideal scenarios(Synthetically generated datasets) and real-world datasets, considering variations in the length and the entropy rate of the data.

Tasks

1. Get acquainted with symbolic time series entropy rate estimation and replicate the results in [3]
2. Implement at least two naive algorithms for entropy rate estimation of time series [1, 2].

3. Evaluate the performance of the implemented algorithms in terms of effectiveness and efficiency for various data sets.
4. Optional: improve the effectiveness and/or efficiency of the estimators.
5. Write the thesis.
6. Optional: Present the thesis work in a DBTG meeting

References

[1] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 2nd edition edition, 07 2006.

[2] Kontoyiannis I., Algoet P. H., Suhov Yu. M., and Wyner A. J. Nonparametric entropy estimation for stationary processes and random fields, with applications to english text. *IEEE Transactions on Information Theory*, 44:1319–1327, 1998.

[3] Annick Lesne, Jean-Luc Blanc, and Laurent Pezard. Entropy estimation of very short symbolic sequences. *Phys. Rev. E*, 79:046208, Apr 2009.

Supervisors: Sven Helmer (helmer@ifi.uzh.ch), Jamal Mohammed (mjamal@ifi.uzh.ch)

Start date: 1 January 2024
End date: 30 June 2024

University of Zurich
Department of Informatics

Prof. Dr. Michael Böhlen
Professor