ulm university universität ulm

Bachelor / Master Thesis

Susceptible Artificial Data Generation

Context

Synthetical data generation using several methods (classic statistics, GANs) is necessary in the context of data analysis for several purposes such as simulation, evaluation and anonymization.

Hereby, not only a synthetization in terms of similar data reproduction with an obfuscation effect is of interest, but also data generation in a “what-if” scenario. In other words, synthetic data generation based on input data and additional parameters (i.e. how would the data look with more users?). Another example is the supplementation of given datasets with anomalies using generative networks.

Such data can then be used not only in evaluation scenarios where the real data might be too sensitive, but also find its way into prediction applications, where the validation can be performed using such data.

Scope

In the scope of this work, an approach for generating synthetic data shall be developed, with the possibility of adding a user defined bias in the process of the generation. A fitting validation approach shall be implemented for the evaluation of the diffused data.

Requirements and Comments

If this thesis achieves the desirable outcome, its results are to be integrated in the RECAP research project (https://recap-project.eu/) which is released under an OpenSource license. For that reason, we appreciate if you are ready to OpenSource your results.

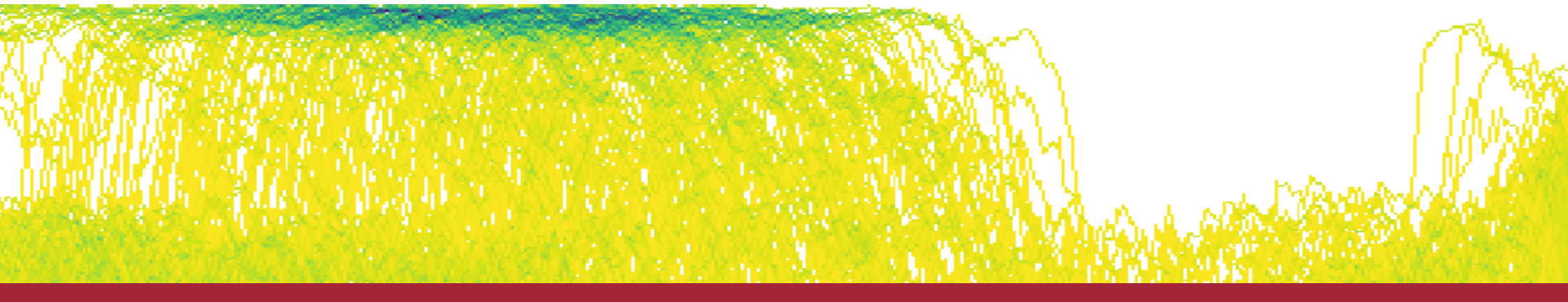
If you are interested in this or similar theses, please contact Mark Leznik either by mail or directly in his office.

mail: mark.leznik@uni-ulm.de
office: Uni West, 43.2.213

Faculty of Engineering and Computer Science

Institute of Information Resource Management

ulm university universität ulm



Bachelor / Master Thesis

Visual Analysis of Time Series Data

Context

Time-series data analysis, specifically anomaly detection (collective anomalies) are still a challenge to recognize for machine learning and classical statistical approaches. The use of CNN and similar deep learning-based classification methods has proven to be successful in image based classification. In this work, the possibility of using classification techniques to recognize anomalies based on a labeled dataset shall be explored.

Scope

In the scope of this work, based on the creation of so-called density maps (optionally any other more suitable visualization), a classification approach for recognizing time series anomalies shall be developed. The validation of the work is to be performed using an open source labeled time series dataset.

Requirements and Comments

If this thesis achieves the desirable outcome, its results are to be integrated in the RECAP research project (https://recap-project.eu/) which is released under an OpenSource license. For that reason, we appreciate if you are ready to OpenSource your results.

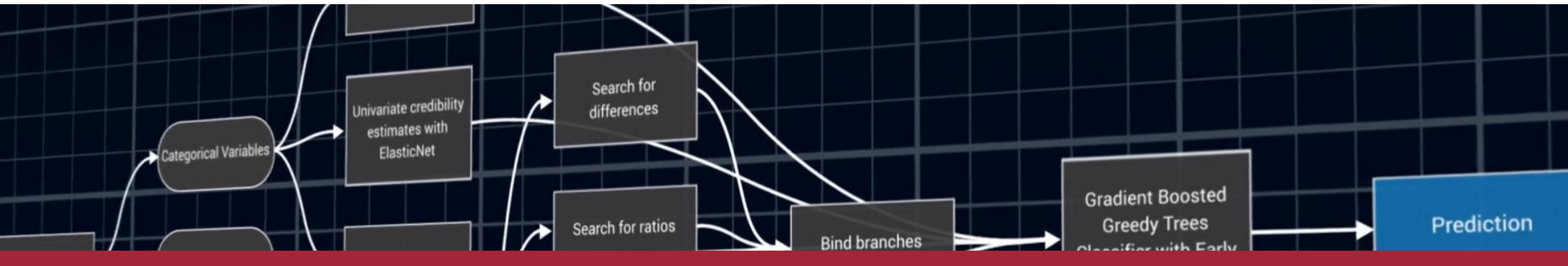
If you are interested in this or similar theses, please contact Mark Leznik either by mail or directly in his office.

mail: mark.leznik@uni-ulm.de
office: Uni West, 43.2.213

Faculty of Engineering and Computer Science

Institute of Information Resource Management

ulm university universität ulm



Bachelor / Master Thesis

AutoML Time Series Analysis

Context

Auto machine learning approaches for data analysis are a promising up and coming research field for several proposes ranging from image analysis to speech recognition. Auto machine learning algorithms based on the minimum nescience principle have also shown to be successful for time series prediction and classification. The algorithms include Decision Trees, Multilayer Perceptrons, and Grammatical Evolution.

Scope

In the scope of this work, the possibility of using an AutoML framework for time series analysis shall be explored. A comparison to more established approaches during the evaluation is order hereby.

Requirements and Comments

If this thesis achieves the desirable outcome, its results are to be integrated in the RECAP research project (https://recap-project.eu/) which is released under an OpenSource license. For that reason, we appreciate if you are ready to OpenSource your results.

If you are interested in this or similar theses, please contact Mark Leznik either by mail or directly in his office.

mail: mark.leznik@uni-ulm.de
office: Uni West, 43.2.213

Faculty of Engineering and Computer Science

Institute of Information Resource Management

ulm university universität ulm

Bachelor / Master Thesis

General Game Playing Approaches for Financial Time Series Analysis

Context

General game playing (GGP) algorithms are designed to by applicable to more than one game successfully. The release of AlphaZero by Deepmind (Google), has led to an increased popularity of this topic and the research around it. Techniques of applying algorithms to a partially known problem with imperfect information, like a game of No Limit Poker, where in comparison to chess or Go, the players do not have the same information, have also emerged. The analysis is financial time series data seems (on the surface at least) as a suitable candidate of applying such approaches to it, due to its imperfect information and large amount of possible outcomes.

Scope

In the scope of this work, the possibility of using a GGP approach and applying it to the process of financial time series analysis shall be evaluated. For this purpose, a survey of the recent state of the art around the subject needs to be conducted. Further, an applicability of such an algorithm on this specific problem shall be evaluated (e.g. in the form of a prototype).

Requirements and Comments

If this thesis achieves the desirable outcome, its results are to be integrated in the ALTO research project (Eurostars ALTO/AII) and released under an OpenSource license. For that reason, we appreciate if you are ready to OpenSource your results.

If you are interested in this or similar theses, please contact Mark Leznik either by mail or directly in his office.

mail: mark.leznik@uni-ulm.de
office: Uni West, 43.2.213

Faculty of Engineering and Computer Science

Institute of Information Resource Management