**FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE**

Review report of a final thesis

Reviewer:	Mgr. Petr Novák, Ph.D.
Student:	Bc. Eliáš El Frem
Thesis title:	Market Data Anomaly Detection
Branch / specialization:	Knowledge Engineering
Created on:	2 June 2024

Evaluation criteria

1. Fulfillment of the assignment

[1] assignment fulfilled
▸ [2] **assignment fulfilled with minor objections**
[3] assignment fulfilled with major objections
[4] assignment not fulfilled

Most of the parts of the assignment were completed, but some only very briefly. Experiments with hyperparameter settings were not conducted.

2. Main written part   65 /100 (D)

The thesis is written in coherent English and is well arranged into chapters. The quality of the text is sometimes brought down by misspellings and typographical errors.

The review of available methods is comprehensive, but some methods could be explained in more detail. In particular, the description of the decision rules or thresholds, how specifically each method identifies the anomalies, is missing.

The description of obtaining the dataset is very brief, mentioning only a website. It is unclear whether an API was used or if the data was obtained manually, and whether any additional preprocessing was necessary. However, the chosen market data time series are interesting and well described.

The main contribution of the work is in the experimental part. However the results are not always easy to read, due to the considerable number of combinations of examined assets, data transformations, and methods used. Numerous graphs are presented, with their descriptions sometimes several pages away. It would be more coherent to focus on one asset class at a time, test all methods on it, and discuss which one works best, then move to another. The evaluation of anomaly detection performance is conducted mostly visually for a large part of the experiments, making it largely subjective. Only in the last few experiments exact metrics are applied. The silhouette score is defined for individual

observations in section 2.2.2, but in section 4.10 it is used to evaluate the entire model. The simulation study on artificial data is interesting but very brief. It would be interesting to explore more situations with varying length of the series and different percentages of anomalies. Evaluation in the form of confusion matrices is not very clear, it would be better to include for example the F1 score or another suitable metric of classifier performance.

The conclusion is more of a summary of the experiments. It lacks an evaluation of the overall contribution of the thesis and possibilities for future work.

3. Non-written part, attachments   50 /100 (E)

The student tested machine learning algorithms for anomaly detection in time series on both real and simulated data. Existing libraries and own implementations in Python are used.

The code is included in a single Jupyter notebook, which is only scarcely commented. The data is not available, so a large part of the experiments cannot be replicated.

4. Evaluation of results, publication outputs and awards   70 /100 (C)

The results of the experiments on market data are interesting and can be used in practice to choose a suitable method for anomaly detection.

The overall evaluation   60 /100 (D)

The applicant studied the problem of anomaly detection in market data time series, he described and tested several methods and performed an extensive array of experiments. The methods could be explained in more depth, the experiments on real data could be presented in a clearer form and the simulation study could be expanded to include more situations and a more precise evaluation. Despite these shortcomings, I believe that the work meets the requirements for a diploma thesis and recommend it to be recognized as such.

Questions for the defense

What thresholds or rules do the individual methods use for identifying anomalies?

How could be explained the inability to replicate the method from the literature in section 4.7?

How are the overall silhouette scores calculated in table 4.1?

Instructions

Fulfillment of the assignment

Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Main written part

Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies?

Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 52/2021, Art. 3.

Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Non-written part, attachments

Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Evaluation of results, publication outputs and awards

Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

The overall evaluation

Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.