CZECH TECHNICAL UNIVERSITY IN PRAGUE
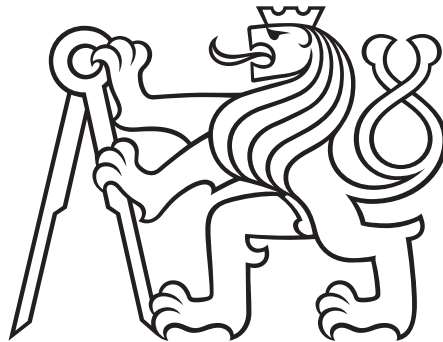
FACULTY OF MECHANICAL ENGINEERING

DOCTORAL THESIS STATEMENT

Czech Technical University in Prague

Faculty of Mechanical Engineering

Department of Instrumentation and Control Engineering

Summary of Dissertation thesis

Novelty detection via linear adaptive filters

Ing. Matouš Cejnek

Doctoral Study Programme: Mechanical Engineering

Branch of study: Control and Systems Engineering

Supervisor: *Doc. Ing. Ivo Bukovský, Ph.D.*

Dissertation thesis statement presented in candidacy for the academic
title of "Doctor" abbreviated to "Ph.D."

Prague   2020

Title in Czech language: *Detekce novosti pomocí lineárních adativních filtrů*

This doctoral thesis is an outcome of a part-time doctoral study programme at the Department of Instrumentation and Control Engineering, Faculty of Mechanical Engineering, Czech Technical University in Prague.

Disertant: *Ing. Matouš Cejnek*
Department of Instrumentation and Control Engineering, Faculty of Mechanical Engineering, Czech Technical University in Prague

Supervisor: *Doc. Ing. Ivo Bukovský, Ph.D.*
Department of Mechanics, Biomechanics and Mechatronics, Faculty of Mechanical Engineering, Czech Technical University in Prague

Reviewers:

The thesis was set out on:

The defense of the dissertation thesis will take place on:

The thesis is available in the Department of Science and Research of Faculty of Mechanical Engineering, CTU in Prague, Technická 4, Praha 6 -Dejvice.

prof. Ing. Tomáš Vyhlídal, Ph.D.
Head of Doctoral Study Field Control and Systems Engineering Faculty of Mechanical Engineering CTU in Prague

Abstract

Novelty detection is an important signal processing task. This task is essential for many industry, and biomedical applications. This thesis is presenting research on the topic of novelty detection utilizing parameters of linear adaptive filters. A new method of adaptive novelty detection is presented in this thesis - Error and Learning Based Novelty Detection. The goal of this thesis is to present the new method as a viable tool for online unsupervised novelty detection in non-stationary and drifted data. The method is supported with various experimental evidence collected from multiple studies. These studies cover multiple traditional applications like system change point detection and outlier detection. The results are obtained from experiments with real and synthetic data.

Abstrakt

Detekce novosti je důležitá část zpracování signálů a je esenciální pro různé průmyslové a bioinženýrské aplikace. Tato disertace prezentuje výzkum metod detekce novosti využívající parametry adaptivních filtrů. V této práci je popsána nová methoda adaptivní detekce novosti nazvaná Error and Learning Based Novelty Detection. Cílem této práce je popsat tuto novou metodu jako užitečný nástroj pro online detekci novosti pro data, která jsou nestacionární nebo obsahují drift. Studie podporující užitečnost této metody jsou představeny v této práci. Tyto studie pokrývají různé oblasti tradičního zpracování signálů, například : detekce změny chování systému a detekce anomálií. Experimentální výsledky těchto studií jsou získány na reálných i syntetických datech.

Contents

1 Introduction

1.1 What novelty detection is

Novelty detection is the name for identification of something new or unknown in data. The exact meaning of event described as something new depends on its application and on the field. However, in general the novel event is something that is not expected in the data because of the data generating process nature.

The task of novelty detection is one of the oldest and the most fundamental tasks in *machine learning* field. The monitoring of production process or of any other process is a costly work if it is done by a human operator. That is the reason why huge effort to automate this process has been done in the last few decades. Despite of this fact, the term novelty detection has started appearing in literature after year 2000. In some cases the goal of novelty detection can be perturbation detection (one-sample-outliers) in gradually changing environment. In other case, the goal may be detection of gradual changes of environment by itself while the perturbations are ignored. Various categorizations of novelty detection are often used. Probably the most fundamental categorization is based on the scale of the detection [7]. The categories are:

- contextual novelty detection - can be understand as system change point detection, or detection of change in process that is generating the data;

- value based novelty detection - this name stands for detection of various perturbations, or generally short-time events, that does not belong into expected behavior of the observed system.

1.2 Adaptive novelty detection

Adaptive novelty detection is a special case of novelty detection. It is a special case because it is featuring adaptive or learning algorithms. An interesting feature of adaptive models is their prediction or classification error. Such an error can provide valuable information about novelty hidden in data.

Summary of the key features of adaptive models that are desirable during novelty detection process is as follows:

- *Compression* - the adaptive algorithms has the ability to describe a long window of historic information in smaller number of parameters of their updates.

- *Prediction* - the ability to predict few samples ahead can be useful to minimize the delay between sample acquisition and evaluation of sample novelty.

- *Compensation* - the adaptation by itself is a mechanism how to compensate for gradual changes in data

Because of the reasons mentioned above, the adaptive novelty detection is a promising field of the machine learning for future research.

1.3 Novelty detection implementation challenges

In general, a novelty detection process can by understood as a type of classification. However, the type of the classification used for the novelty detection varies according to the used approach. Some methods dealt with the novelty detection as with a binary classification - the first class is a normal event and the second class is a novel event. Other methods used multiple classes, where one or more classes represents the novel events. Also it is not uncommon to understand the novelty detection as a classification with unknown class or classes.

The issue of novelty detection as a classification is evident. A sufficient training sample is a must for any successful classification. However novelty detection is a task commonly defined without any information about how the novel events should look like. Basically the novelty detection is search for something that is not known, hard to model and difficult to predict.

However, in specific cases it is possible to at least annotate the novel event retrospectively and thus the novelty detector (classifier) can be supervised, or at least some kind of *reinforcement learning*[1] can be applied.

---

[1]Reinforcement learning is process where some kind of reward function is used instead of an exact input/output pair as a feedback.

2 The state of the art

2.1 Novelty detection concepts introduction

The novelty detection may be supervised and also unsupervised. Supervised means, that we have some information how the novelty in the data should look and thus we can train a model to search for this novel events or objects. The term unsupervised novelty detection means, that we do not know what is the novelty in given data-set and we need some method to identify and describe those not common pieces of data. Note that the supervised novelty detection is not too different from an ordinary two-class classification. The classification approach is suitable for any cases, where it is possible to measure whether the finding of the novel event was correct.

Novelty detection methods are commonly separated into statistical based and neural network based. Also it is common to combine something from both approaches in one method. A statistical approach of novelty detection uses statistical properties of the acquired data to decide, whether the data is novel or not. Statistical novelty detection methods could be divided between two groups - parametric approaches and non-parametric approaches. Parametric approaches expect, that distribution of evaluated data is Gaussian in nature. It means, that the data distribution can be modelled just with the data mean and covariance. Non-parametric approaches are more flexible, because they do not have assumptions about the data distribution form. This cause that they are also more computationally expensive [23] in general. However, it is important to note that literature on this topic is not completely united in opinion what methods are parametric and what methods are non-parametric [11].

The second category of novelty detection methods is based on artificial neural networks (ANN) [34]. Such methods are heavily used for novelty detection tasks today, because of the recent general popularity of ANN. The ANN based methods have advantages and also disadvantages in comparison to the statistical approaches. Probably the main advantage of the ANN based methods is possibility of online retraining. Commonly discussed disadvantage of the ANN is the huge dependence on the chosen ANN architecture

and the complexity of its optimization [24]. If the chosen ANN architecture is too simple, it may have difficulties to learn the system properly. On the other hand, if the architecture is too complicated, it may lose the ability of the generalization that leads to bad performance. For selection of a correct ANN architecture a few approaches exist. The most common and intuitive ones are performance testing while pruning - decreasing complexity, and performance testing while increasing the architecture complexity (also known as constructive algorithms). In general the most common ANN architecture also for novelty detection is multi-layer perceptron (MLP) [35]. The confidence measure of a MLP input patterns is popular novelty indicator. The simplest method how to achieve that is to put a threshold on the ANN output. In other words, the MLP recognize whether the new pattern is know or unknown.

2.2 Main approaches to novelty detection

2.2.1 Hypothesis testing

Hypothesis testing belongs into the group of parametric approaches. It is simple statistical method commonly used for testing whether the tested data or sample belongs to the same distribution as training data or not. The test popular for this topic is the *Grubbs' test* [13]. The *Grubbs' test* is based on comparison of distance from the test data points and the sample mean. Any data point with this distance higher than a certain threshold is considered as an outlier. The popular value for the threshold is commonly the value of three standard deviations from the mean value. This test assumes that the training data posses Gaussian distribution and it works only with univariate continuous data. Many variants of this test was proposed later to deal with multivariate data sets, for example [12].

2.2.2 Gaussian mixture model

Gaussian mixture models (GMM) [10] is a parametric probabilistic approach. This approach is based on the idea that data are generated from a weighted mixture of Gaussian distributions. Thus it is considered as a parametric approach. However, in practice the GMM approach (or similar) has a problem with the dimensionality of data. With the high dimensionality, this method needs a very large number of samples to train a model [23]. Another problem is the correct selection of suitable threshold.

2.2.3 Hidden Markov models

Hidden Markov models (HMM) [9] are stochastic models for sequential data, and belong into group of parametric approaches. The HMM is build on assumption, that modeled system is process with unobserved hidden states - Markov process. The transition from the hidden states to observable states is done via stochastic process. Every observable state is associated with a set of probability distributions. The change in probabilities of any observable event is compared to a threshold to test for novelty. The hidden Markov models are generally popular for pattern recognition, for example: temporal pattern recognition in bioninformatics, speech, gesture recognition, handwriting recognition and similar tasks. The issue of this method is that modeled distribution must be Gaussian in nature.

2.2.4 Support vector machines based approach

Support Vector Machines (SVM) [36] is a not-probabilistic method [19]. It is designed to work as a binary classifier. In other words the SVM algorithm search for best hyperplane to separate the known classes. Although the algorithm was originally created as linear classifier, it was later extended with kernel transformation to nonlinear classifier. The kernel transformation can change linearly inseparable task into the task that is linearly separable.

2.2.5 K-nearest neighbour algorithm

K-nearest neighbour algorithm (KNN) [37] is non-parametric statistical method used for classification and regression. The KNN based methods are among the most popular methods for novelty detection nowadays [11]. The KNN algorithm assigns the class for every new sample according to the class of k nearest neighbours. The KNN novelty detection approach is based on the assumption that normal data-points have close neighbours, while novel points are located far from those points [8]. However, the KNN technique has problem with huge data-sets, because its evaluation demands much bigger number of computational operations [23].

2.2.6 Neural networks clustering based methods

Clustering is operation of splitting data into groups. For novelty detection it means classification to normal data and novel data. It could be either supervised or unsupervised. Probably the most common ANN for clustering is self-organizing map (SOM) [21]. As the survey [7] suggests, the SOM is also

really popular concept for novelty detection. It is non statistical alternative to clustering algorithms. Although the clustering approaches have shown some potentionals in mentined studies, they struggle with the issues of correct segments selection and feature extraction [22] in general.

2.2.7 Reconstruction based approaches

Reconstruction based approaches are methods based on data reconstruction (current or historical sample estimation from reduced features) or prediction. Various ANN architectures or similar learning models could be used for this purposes. Basic idea is that the adaptive model is trained for reconstruction (prediction) of input data. When input data vary from training data, reconstruction error rises up. This method could be used online with forward reconstruction. Big advantage of this approach is simple retraining (online adaptation). During the retraining the increment of adaptive parameters of the auto-associator could be used as a novelty indicator. These features make the reconstruction based approaches excellent for online processing of data streams. Methods based on data reconstruction seems to work well. But they also have the same issues like other ANN methods - it is hard to make a mathematical evidence why it works [20]. Also the computational demands of such algorithms could be overwhelming for some applications.

2.3 Open problems

Major novelty detection methods were presented in this chapter. Most of these methods feature at least one of the following issues:

- A method needs an a-priori information about the novelty and/or healthy data. In other words, the method works as a classifier for known classes.

- A method needs a heavy pre-training and/or has a great time complexity. Thus the method is suitable only for offline use.

- A method heavily relies on statistical attributes of the data. Therefor the method has a hard time to deal with non-stationary process - data with any kind of concept drift.

3 | Thesis objectives

As it was introduced in previous section, the machine learning field has currently huge demand for algorithms that can work for data streams produced in real time. According to this demand, the first objective of the thesis is set to:

1. objective - *Development of an adaptive novelty detection method suitable for online data streams processing.* Such a method should be able to re-adapt to new data on the fly without the need for any time expensive re-learing.

The optional but often required feature of novelty detection methods for data streams is computation speed. Not every real-time process use high sampling rate, however a low lag novelty detection is generally beneficial. Because of that reason, the second objective of this study is

2. objective - *Development of a fast adaptive novelty detection method applicable with fast adaptive algorithms.* Such a method should have low time complexity. Thus, it should be suitable for machines with low computational performance.

The online data streams processing posses the issue of concept drift and other significant data imbalances that cannot be removed in real-time. Hence the third objective of this thesis is set to:

3. objective - *Development of an adaptive novelty detection method robust against concept drift and non-stationary data.* Data trends and drifts are common and it is hard to deal with them in a real-time processing. Therefore the proposed algorithm should be able to compensate for data drifts on the fly.

4 | Developed method

4.1 Method description

The proposed method is called Error and Learning Based Novelty Detection (ELBND). In this work, various types of adaptive filters are used as the base for adaptive models. An adaptive filter is a system with a linear filter that has a transfer function controlled by variable parameters to adjust those parameters according to an optimization algorithm.

The output of adaptive filter or of any similar adaptive model could be described with the following equation

$$\tilde{y}(k) = w_1(k) \cdot x_1(k) + ... + w_n(k) \cdot x_n(k) = \sum_{i=1}^{n} w_i x_i = \mathbf{w}^T(k) \cdot \mathbf{x}(k), \qquad (4.1)$$

where k is discrete time index, $\tilde{y}(k)$ is output (filtered) signal, $\mathbf{w}$ is vector of adaptive weights, $\mathbf{x}$ is input vector and $(.)^T$ denotes the transposition. The initial values of adaptive weights (adaptive parameters) $\mathbf{w}$ are usually set to all zeros, or alternatively to random numbers (normal distribution, zero mean value). The $\mathbf{x}(k)$ is input vector. The bias should mimic the bias in neural units. In practice, this bias can compensate offsets and similar data imbalances. The mentioned error e of the adaptive filter is calculated as

$$e(k) = y(k) - \tilde{y}(k). \qquad (4.2)$$

The second input of the method is the increment of adaptive weights defined as

$$\Delta \mathbf{w}(k) = \mathbf{w}(k+1) - \mathbf{w}(k). \qquad (4.3)$$

The method of the increment $\Delta \mathbf{w}(k)$ estimation depends on chosen learning algorithm. The proposed way how to combine the parameters $\Delta \mathbf{w}(k)$ and error $e(k)$ to obtain the descriptor of novelty in given sample can be described as follows

$$\mathbf{nd}(k) = \left| e(k) \cdot \Delta \mathbf{w}(k) \right|. \qquad (4.4)$$

The novelty descriptor $\mathbf{nd}(k)$ is vector of coefficients describing how much

novelty is encounter with individual weights $\Delta\mathbf{w}(k)$.

For some applications it could be desirable to describe novelty in data just with single value for every sample. As a good practice how to achieve that, it is reduction of this vector $\mathbf{nd}(k)$ to scalar as follows

$$nd(k) = \max(\mathbf{nd}(k)). \tag{4.5}$$

However, other function than max might be used.

4.1.1 LMS adaptive filter

The classical least means squares algorithm (LMS) [25] is stochastic gradient descent method. It is probably the most common algorithm for adaptive filters. The LMS weights adaptation could be described as follows

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \Delta\mathbf{w}(k), \tag{4.6}$$

where $\Delta\mathbf{w}(k)$ is

$$\Delta\mathbf{w}(k) = \mu \cdot e(k) \cdot \frac{\partial y(k)}{\partial \mathbf{w}(k)} = \mu \cdot e(k) \cdot \mathbf{x}(k), \tag{4.7}$$

where μ is the learning rate (step size) and The general stability criteria of LMS [25] stands as follows

$$|1 - \mu \cdot \mathbf{x}(k)^T \cdot \mathbf{x}(k)| \leq 1. \tag{4.8}$$

The novelty detection could be done as follows

$$\mathbf{nd}(k) = \left| e(k) \cdot \Delta\mathbf{w}(k) \right| = \left| e(k)^2 \cdot \mathbf{x}(k) \cdot \mu \right|. \tag{4.9}$$

4.1.2 NLMS adaptive filter

The normalized least mean squares (NLMS) [25] adaptive filter is extension of LMS adaptive filter. The NLMS adaptation rule could be described as follows

$$\Delta\mathbf{w}(k+1) = \frac{\mu}{\epsilon + \mathbf{x}(k)^T \cdot \mathbf{x}(k)} \cdot \mathbf{x}(k) \cdot \mathbf{w}(k) = \eta \cdot \mathbf{x}(k) \cdot \mathbf{w}(k), \tag{4.10}$$

where ϵ is a constant (regularization term) introduced to preserve stability for inputs close to zero [18].With the NLMS adaptive filter the novelty in data

could be calculated as follows

$$\mathbf{nd}(k) = \left| e(k) \cdot \Delta \mathbf{w}(k) \right| = \left| e(k)^2 \cdot \mathbf{x}(k) \cdot \eta \right| = \left| \frac{e(k)^2 \cdot \mathbf{x}(k) \cdot \mu}{\epsilon + \mathbf{x}(k)^T \cdot \mathbf{x}(k)} \right|. \qquad (4.11)$$

4.1.3 LMF adaptive filter

The least mean fourth algorithm (LMF) [25] is slight modification of the LMS algorithm. The LMF weights adaptation $\Delta \mathbf{w}(k)$ is calculated as follows

$$\Delta \mathbf{w}(k) = \mu \cdot e(k)^3 \cdot \frac{\partial y(k)}{\partial \mathbf{w}(k)} = \mu \cdot e(k)^3 \cdot \mathbf{x}(k), \qquad (4.12)$$

The ELBND is then calculated as follows

$$\mathbf{nd}(k) = \left| e(k) \cdot \Delta \mathbf{w}(k) \right| = \left| e(k)^4 \cdot \mathbf{x}(k) \cdot \mu \right|. \qquad (4.13)$$

According to the (4.13) the ELBND emphasize the error of adaptive filter more with the LMF than the plain LMS.

4.1.4 NLMF adaptive filter

The normalized least mean fourth (NLMF) is often used because it has greater ability to suppress noise than NLMS adaptive filter according to study [4]. On the other hand, it is much harder to enforce stability of the NLMS filter than NLMS filter [5, 6]. The NLMF adaptation [25] is similar to NLMS adaptation. The vector of adaptive weights of a NLMF filter $\mathbf{w}$ is done according to the rule

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \Delta \mathbf{w}(k) = \mathbf{w}(k) + \eta(k)\mathbf{w}(k)e(k)^3, \qquad (4.14)$$

where $\eta(k)$ has the same meaning like in (4.10). With the NLMF adaptive filter the novelty in data could be calculated as follows

$$\mathbf{nd}(k) = \left| e(k) \cdot \Delta \mathbf{w}(k) \right| = \left| e(k)^3 \cdot \mathbf{x}(k) \cdot \eta \right| = \left| \frac{e(k)^3 \cdot \mathbf{x}(k) \cdot \mu}{\epsilon + \mathbf{x}(k)^T \cdot \mathbf{x}(k)} \right|. \qquad (4.15)$$

4.1.5 GNGD adaptive filter

The generalized normalized gradiend descend (GNGD) adaptive filter [18] is an extension of the NLMS adaptive filter. The adaptive weights of a GNGD filter $\mathbf{w}$ are adapted according to the same rule as NLMS. The difference is in parameter $\eta(k)$. The adaptive learning rate (step size) $\eta(k)$ is estimated

in similar way like for NLMS or NLMF, however the regularization term ϵ is obtained in the way that follows

$$\epsilon(k) = \epsilon(k-1) - \rho\mu\frac{e(k) - e(k-1)\mathbf{x}^T(k)\mathbf{x}(k-1)}{(||\mathbf{x}(k-1)||^2 + \epsilon(k-1))^2},\tag{4.16}$$

where the ρ is a custom parameter. As proposed in [18] the GNGD the method should be robust if the parameter ρ is set to small (< 1) positive number. The resulting GNGD formula for novelty detection can be combined from (4.16) and (4.11).

4.1.6   RLS adaptive filter

Other algorithm what could be used for the novelty detection is Recursive least squares (RLS) [25]. For this method the adaptive weights are calculated as follows

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{R}^{-1}(k)\mathbf{x}(k)e(k)\tag{4.17}$$

where the matrix $\mathbf{R}^{-1}(k)$ is inverse of the auto-correlation matrix with size $n \times n$, where n is number of adaptive weights $\mathbf{w}(k)$. The $\mathbf{R}^{-1}(k)$ matrix is obtained as follows

$$\mathbf{R}^{-1}(k) = \frac{1}{\gamma}\left(\mathbf{R}^{-1}(k-1) - \frac{\mathbf{R}^{-1}(k-1)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{R}^{-1}(k-1)}{\gamma + \mathbf{x}^T(k)\mathbf{R}^{-1}(k-1)\mathbf{x}(k)}\right).\tag{4.18}$$

According to the adaptation rule (4.17), we can describe proposed novelty detection method with RLS algorithm as follows

$$\mathbf{nd}(k) = \left|e(k) \cdot \Delta\mathbf{w}(k)\right| = \left|e^2(k)\mathbf{R}^{-1}(k)\mathbf{x}(k)\right|.\tag{4.19}$$

5 | Results

5.1 Nonstationary biomedical data

Biomedical sciences are a set of applied sciences derived from natural science dealing with healthcare or public health. The data processing requests by biomedical science researches are mainly related to signals measured on human body. These signals are naturally complex and non-stationary. Here are introduced two studies of the proposed novelty detection method use for biomedical data processing - the detection of perturbations in ECG and the classification of Alzheimer's disease from EEG signal. These two studies should demonstrate the abilities of ELBND to deal with non-stationary and offseted real-time signal.

5.1.1 Perturbation detection in ECG

This study was presented in [17]. The goal of this study was detection of artificial perturbations in ECG signal. The ECG measurement is a process of recording the electrical activity of the heart over a period of time using electrodes placed on the skin. A various perturbations can occur in this signal naturally as the ECG artifacts. Such perturbations might decrease the accuracy of further ECG processing or even make it completely impossible. In cases where the perturbations are small, it can be difficult to remove them with conventional methods or by human operator. Because of this reason, the novelty detection can be a simple way how to detect such artifacts in ECG and report them for further processing. The real measured ECG signal was obtained from the Yoshizawa-Sugita Lab (formerly Yoshizawa-Homma Lab), Tohoku university. The used time series was measured by an internal cardio-defibrillator with frequency of 256Hz. This signal was chosen because it contains spontaneous ventricular tachycardia, which is a rare phenomenon to measure. Figure 5.1 shows which part is the healthy ECG signal and which part represents the arrhythmia. In the introduced novelty detection, it is possible to detect the start of the arrhythmia signal, approximately 1000 samples before the arrhythmia is introduced. The included perturbations are not clearly seen in this graph. However, these perturbations are located in the
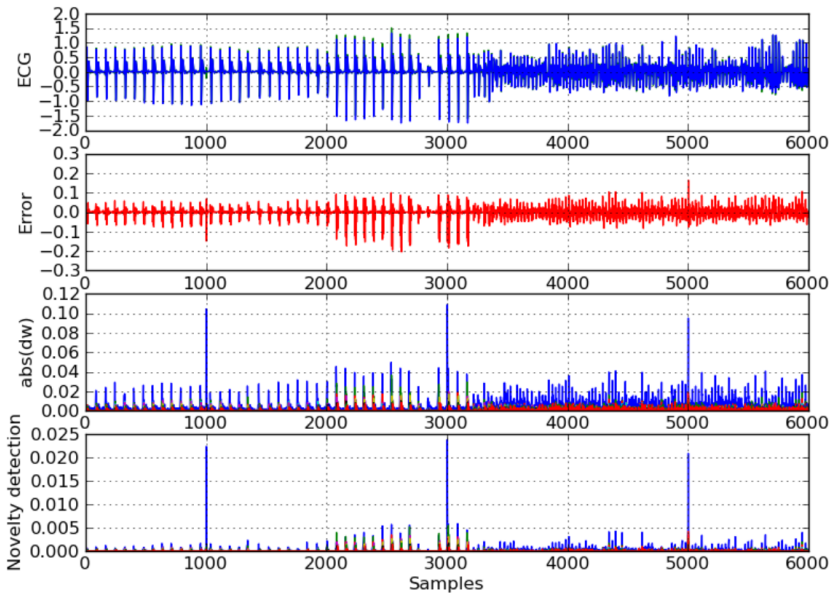
Figure 5.1 – Novelty Detection used on real measured ECG signal. The plot is adopted from study [17].

samples of discrete time 1000, 3000, 5000. On the graph of the prediction error, it is possible to see the perturbations, and even more so in the graph of absolute values of adaptive weights increments. However, looking on the graph of novelty detection, these perturbations are even more evidently pronounced. This experiments demonstrate that the proposed novelty detection method is capable to highlight perturbations in the ECG data, even if the data are contaminated with noise. The simulation was performed on a personal computer and measured speed was higher than what is sampling of ECG signal. In other words, this implementation can be used in real time.

5.1.2 Alzheimer's disease classification

This section is based on study [15]. In this case the novelty detection method was used for extraction of features from EEG signal. The features were used for classification of patients for Alzheimer's disease detection. Data selection contains records from 220 anonymous patients. From that selection, 110 patients match the clinical criteria of dementia and the rest are normal. Every

patient has 2 to 5 manually selected records with length 90s or less.

The EEG signal history was used for adaptation of LNU predictor. The tested history windows for prediction of one sample were really short (4 samples back and 9 samples back). This history cannot contain complete information about signal dynamics. That cause significant prediction inaccuracy. But such inaccuracy is not an issue for this method. The ELBND values were obtained from attributes of predictive model (prediction error, increment of adaptive weights) for every EEG electrode for every patient. Two statistical functions applied on estimated ND were used to created criteria to decide whether the EEG records belong to healthy person of patient with dementia. First investigated function was standard deviation and the second one was entropy.

For validation of the method, non-exhaustive cross-validation was used. The criteria based on standard deviation was just median of all patients from the training group. The entropy based criteria was build on the finding that demented patients have much bigger dispersion of novelty detection entropy than healthy patients. The last criteria uses both statistical functions. The main part of this criteria is standard deviation. This value was further modified with a penalization for entropy. This combined criteria has best classification results. The sensitivity and specificity are both 90%. Dispersion is shown in Fig. **??**.Results of all criteria are summarized in Table 5.1.

Criteria based on	Specificity	Sensitivity
Standard deviation	88%	88%
Entropy	65%	82%
Standard deviation and entropy	90%	90%

Table 5.1 – Table of results for classification based on ELBND method

5.2   Dealing with concept drift

The potentials of the proposed ELBND method for novelty detection with drifted real-time data are presented in this section. This potentials were tested in two studies [1] and [2].

5.2.1   Modeling of concept drift

In a field of novelty detection, the concept drift is considered as a challenging data imbalance that should be ignored, and only system changes and outliers

that represent novelty should by highlighted by the novelty detection methods. The examples of gradual, recurring and abrupt drift are shown in Figure 5.2.

Two types of gradual concept drift were simulated in the studies presented in the next subsections. Two models of concept drift were used in mentioned works:

- Ramp (pure gradual drift - slow constant increment). This kind of drift is problematic because greater distance from data zero mean can alter the adaptive model performance and also it makes any standardization (z-score) impossible.

- Harmonic wave (periodically repeating drift). This drift is also known as recurring trend. This drift is commonly present in various biological systems and stock market behavior.

5.2.2 Reference methods and signals

Error of prediction
The simplest reference signal is just the plain error of the adaptive model. This reference is interesting because it is the most easiest feature describing novelty in data that is possible to obtain from adaptive model. Usage of the error as a reference directly displays how much information can be obtained from the adaptive model with more sophisticated methods like ELBND or LE. That is the reason why it is used in study [1] for result validation.

Learning entropy (LE)
The learning entropy (LE) is more advanced but similar learning-based method to ELBND. That is the reason why it is useful to compare it with ELBND in this thesis. The method called the LE is also called the approximate individual sample learning entropy (see AISLE in [29]).

Sample Entropy
Sample entropy (SE) is a modification of approximate entropy (ApEn), used for assessing the complexity of time-series signals, mainly used for physiological time-series and diagnosing diseased states [26]. The SE is a proven to be a conventional tool for detection of novelty events. It was studied for detection of epilepsy in clinical applications [31]. In [30], neonatal sepsis detection from abnormal heart rate characteristics was studied.
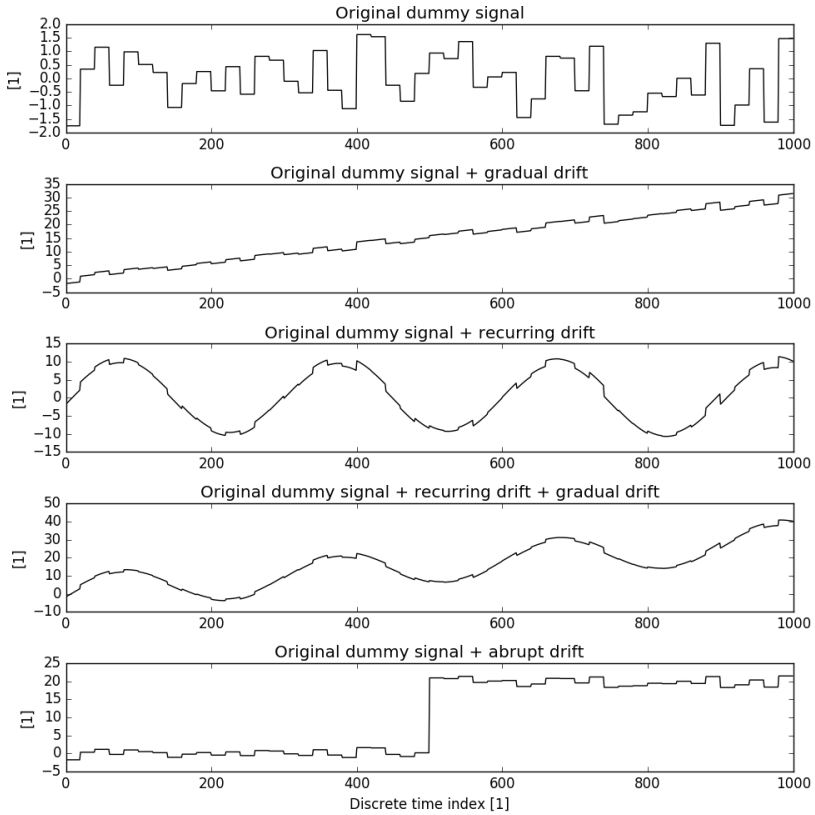
Figure 5.2 – Examples of concept drift effect on synthetic dummy signal.

The SE was chosen as a benchmark tool for the study [1] because it annotates the samples in similar way like the proposed method ELBND, or the similar adaptive method LE. That is the reason why it is interesting for direct comparison also in this thesis.

5.2.3 Experiments and results

Data for all simulation were created synthetically to achieve uniform occurrence of novel events among data. The detailed reasons for this solution were explained in introduction. Two different novelty detection tasks were created for the experimental analysis. First task is the system change point

detection (contextual novelty detection) with a system that can be completely modeled by a used adaptive model. Second task is the outlier detection (value based novelty detection) in a more complex signal (ECG waveform) that is not possible to fully model with a given adaptive model. These two tasks were selected because every one of them represents different challenge for adaptive models and novelty detectors. The experiments are explained in detail in following subsections. All simulations have been done in language Python.
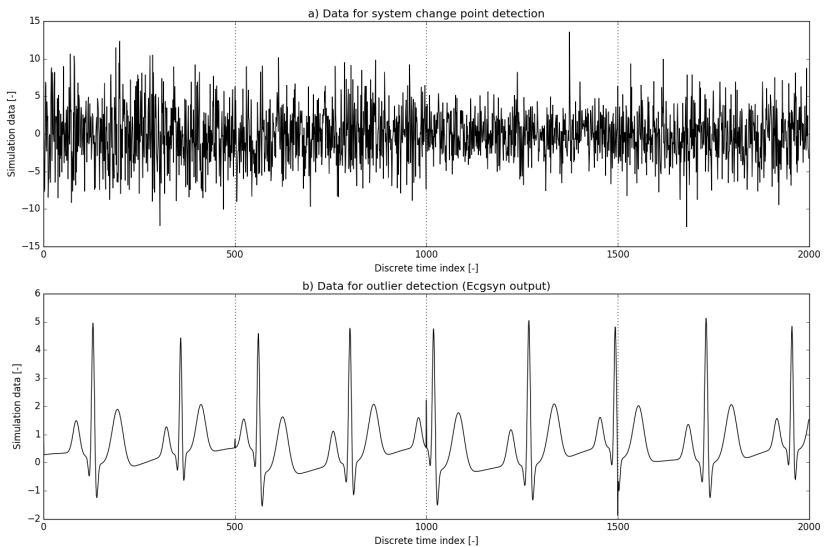

Figure 5.3 – Data used for detection and validation (each of 250 000 samples in total). Doted vertical lines mark the positions of novelty occurrences (of random magnitudes): a) the detail of first data set is the output of system with system changes as novelty; b) the part of second data set for outlier detection - Ecgsyn generated ECG (waveform with perturbations - outliers)

System change point detection with NLMS

This experiment can be found in study [1]. The goal of this experiment was to test the ability of ELBND, LE and SE to detect novel events (system change point) in data. The results of the experimental analysis is interesting for this thesis, because it provides some new information on ELBND novelty

detection potentials. Especially it provides comparison between adaptive based methods (ELBND, LE and error of prediction) and the conventional method SE.

In order to achieve a general data-set for testing novelty detection, first, a system that is possible to be modeled by an adaptive filter with zero error was created. In other words, the model should be able to recognize novel event in data in all cases without a mistake. This ideal system was contaminated with noise and concept drift (5.1) to make the task of novelty detection difficult. With this setup it was possible to monitor how difficult is the environment for the adaptive model.

The used data $y(k)$ were generated according to the following equation

$$y(k) = h_1(k)x_1(k) + ... + h_n(k)x_n(k) + \xi(k) + \chi(k), \tag{5.1}$$

where $h_i(k)$ are parameters of the data generator, $x_i(k)$ are input variables, $\xi(k)$ is noise and $\chi(k)$ represents the concept drift. Ten independent series of white Gaussian noise with unit standard deviation and zero mean were used as the input. The generator parameters $h_i(k)$ change randomly every $n_{change} = 500$ samples. These changes of parameters are sharp and with unit standard deviation and zero mean. The data contains 500 of such changes. The example of the resulting data can be seen at Figure 5.3a. The level of noise for this simulation was 10.429dB on average. Note that the level of noise was slightly different in every segment of data. This variation is caused by different parameters $h_i(k)$ of generator.

The adaptive filter was used in predictive settings with $n = 10$ adaptive parameters (a parameter for an input). At the beginning, the parameters were set to zeros. Initial value for adaptive learning rate was set to $\eta(k) = 1.5$. The resulting receiver operator curve (ROC) are shown in Fig. 5.4.

Outlier detection with NLMS

The goal of this analysis was to correctly detect occurrence of perturbations in data. Synthetic electrocardiography (ECG) data was used for this study. To generate synthetic ECG data, the Ecgsyn [32] (a realistic ECG waveform generator) was used.

Perturbations (simulated outliers) were introduced into this simulated waveform time-series. The outliers were random numbers from normal distribution with zero mean and 0.1 standard deviation (standard deviation of non
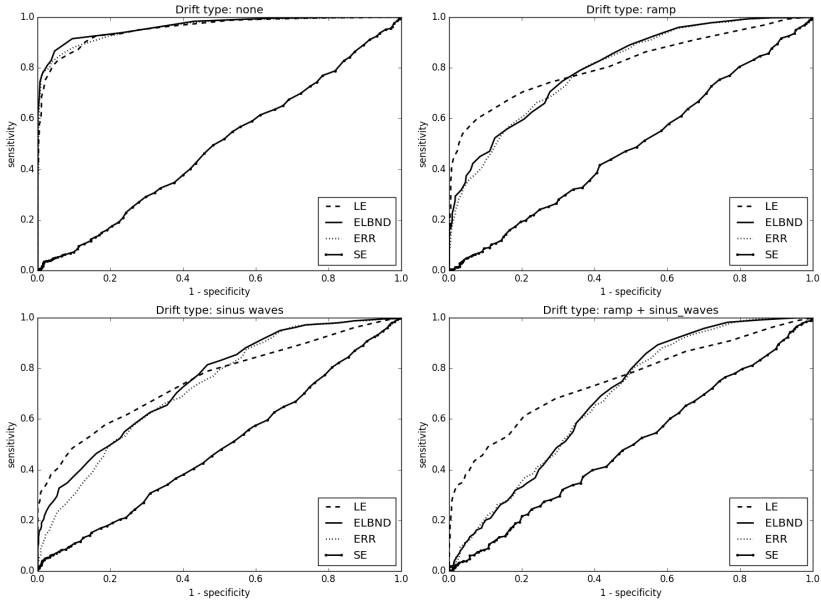
Figure 5.4 – The ROC curves for system change point detection analysis (ERR - error of prediction, SE - based on sample entropy).

drifted ECG signal is 0.907). These outliers were added to the signal values. One outlier was placed at every 500 samples. Total number of introduced outliers was 500. The example of the resulting data can be seen at Figure 5.3b. The resulting receiver operator curves (ROC) are shown in Fig. 5.5.

The adaptive filter was used in predictive settings with $n = 10$ adaptive parameters (a parameter for an input). At the beginning, the parameters were set to zeros. Initial value for adaptive learning rate was set to $\eta(k) = 1.5$. Note that the average period of one ECG wave is about 25.6 times greater than history used for prediction ($n = 10$). Thus, in this experiment the adaptive model cannot fully learn the dynamic behind the ECG generating process. In other words, the model will always have some prediction error, no matter how long it will learn.
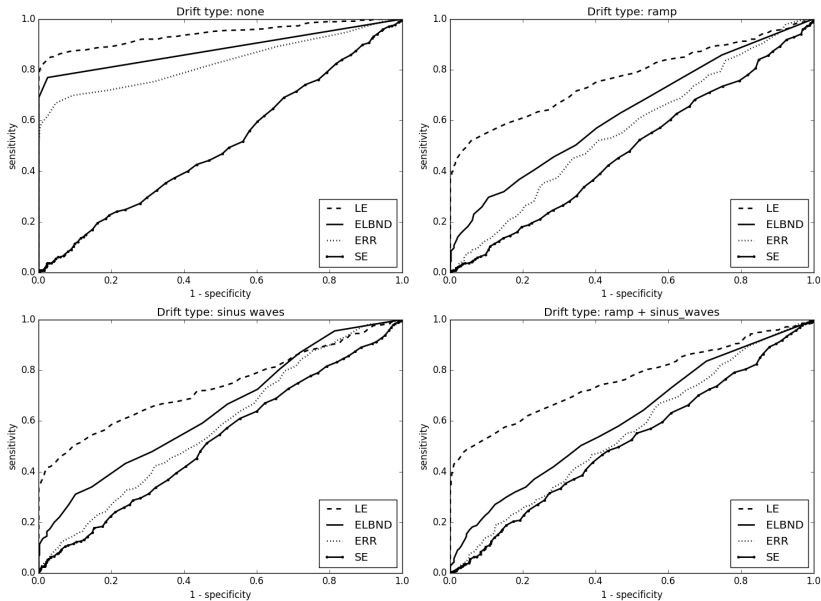
Figure 5.5 – The ROC curves for outlier detection analysis (ERR - novelty detection based on error of adaptive model, data Fig.2b).

Comparison of system change point detection with NLMS, NLMF, RLS and GNGD

The results presented here are obtained from study [2]. Simulated data were used in this study to compare the performance of ELBN and LE in dependency of used adaptive filter (NLMS, NLMF, RLS, GNGD). The main contribution of this study for this thesis is the performance overview of ELBND used with various adaptive filters.

The data used in the study were generated according to the following equation

$$y(k) = h_1(k)x_1(k) + ... + h_n(k)x_n(k), \tag{5.2}$$

where $h_i(k)$ are parameters of the process and $x_i(k)$ represents input variables. There were ten input variables. Each one of them was independent series of white Gaussian noise with unit standard deviation and zero mean value. The process parameters $h_i(k)$ were randomly changed every 500 samples. These sharp changes were introduced to the data as the target novelty.

The actual values of the process parameters $h_i(k)$ were taken from normal distribution with standard deviation of 0.5 and zero mean value. The data-set contain 500 of such sharp changes representing novel events (250×10^3 samples). The generated data were contaminated with *additive white Gaussian noise* (AWGN). At the end, the concept drift was added to the data. The concept drift was modeled as a harmonic wave with period of 10^4 samples and varying amplitude. First two waves (20×10^3 samples) were used for training of adaptive filters. The rest of the data was used for testing.

The process of data generation described above is also shown graphically in Figure 5.6 (drift period=10000, drift amplitude=5, SNR = 10db).
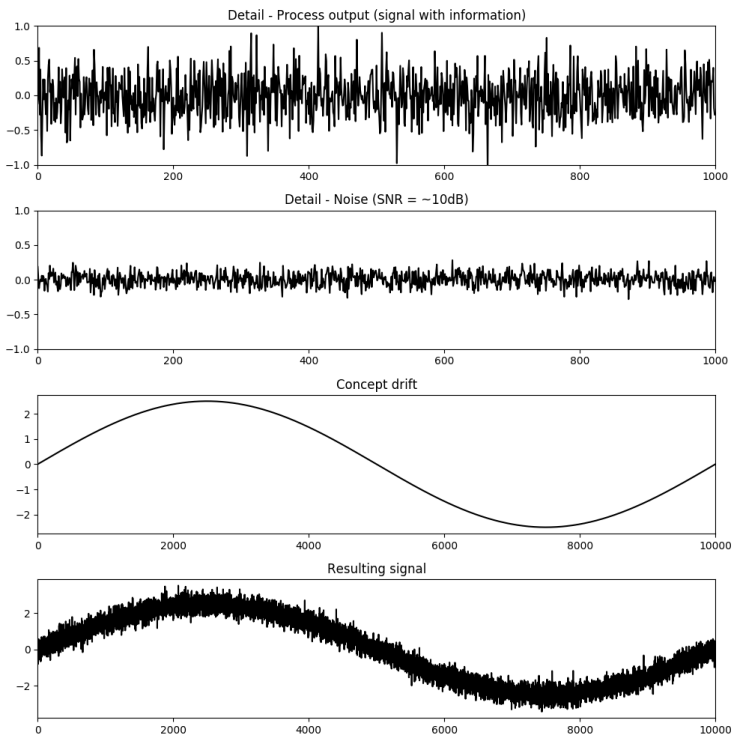

Figure 5.6 – The simulated signal representing output of simulated system and its components.

The results of the experiments were evaluated by three different metrics: AUROC, maximal accuracy (MAX ACC) and ROC (for graphical comparison).

Some of the graphical results - the ROC curves are shown in figures: Figure 5.7, Figure 5.8 and Figure 5.9.
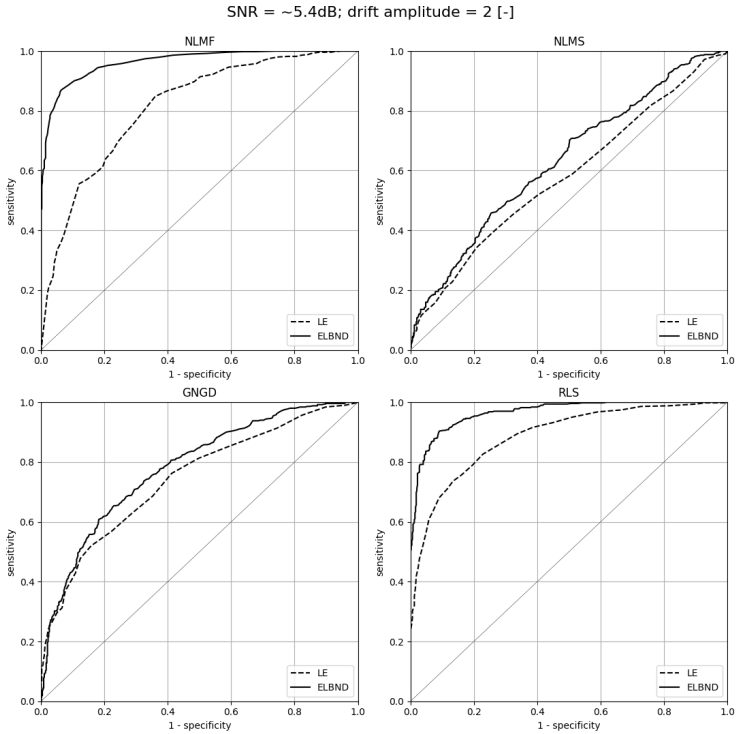

Figure 5.7 – ROC curves for data with concept drift (drift amplitude=2) and with high level of noise (SNR = 5.4dB). In this case ELBND produces better results than LE with all tested adaptive filters. The plot is adopted from study [2].

According to the obtained results, it seems that ELBND has better potential with NLMF and RLS algorithms, while the LE works better with GNGD and NLMS algorithms. It is not surprising because the GNGD and NLMS are very similar gradient methods.

5.2.4 Summary

In this section all results related to the ELBND method used for data with concept drift were presented . Multiple novelty detection tasks were simu-
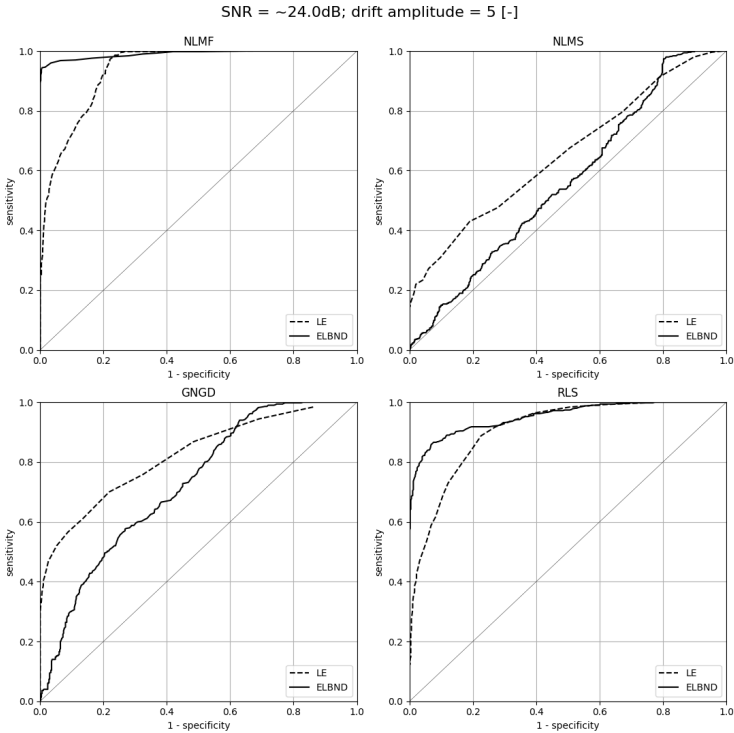
Figure 5.8 – ROC curves for data with drift (drift amplitude=5) and with low level of noise (SNR = 24.0dB). In this case ELBND produces better results than LE only for some adaptive filters. The plot is adopted from study [2].

lated to test the method and their suitability for system change detection and for outlier detection under the occurrence of concept drift, which usually complicates the detection as it appears from comparison to merely detection via plain error based detection and sample entropy based detection.

The results from both presented studies displays that ELBND can compete to LE in various cases. Furthermore, the ELBND extracted feature is generally better than just plain prediction error for novelty detection. According to the second study, it seems that the LE works especially well with RLS and NLMF algorithm. The most important finding is, that the ELBND works in all cases better than the SE algorithm.
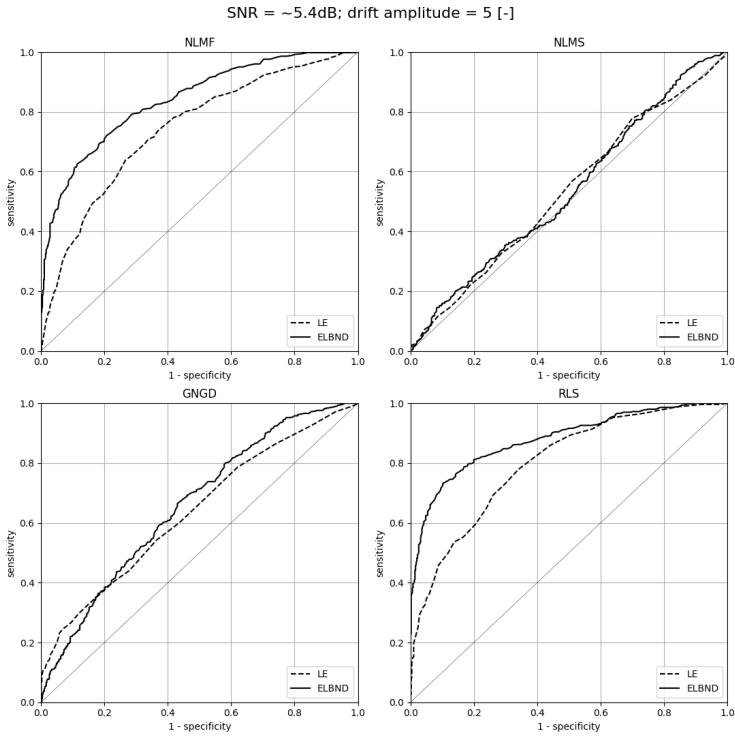
Figure 5.9 – ROC curves for data with drift (drift amplitude=5) and with high level of noise (SNR = 5.4dB). In this case ELBND produces better results than LE only for some adaptive filters. The plot is adopted from study [2].

5.3   Other experiments

5.3.1   ELBND time complexity analysis

This subsection presents results from study [3]. The study investigate time complexity of algorithms for adaptive novelty detection: ELBND, LE, Mahanobilis distance of weights increments (MD) [33] and Fuzzy Density (FD) of weights increments [3].

The time complexity of the ELBND algorithm iteration is broken down step by step in Tab 5.2. As you can see from the table, the complexity of the algorithm strongly relies on the target device, environment and language of

the chosen implementation.

order	operation	complexity	additions	multiplications	note
1.	$o_1 = \Delta \mathbf{w}(\mathbf{k})e$	$O(n)$	0	n	-
2.	$o_2 = |o_1|$	$O(n)$	0	0	abs()
3.	$\max(o_2)$	$O(n)$	0	0	max()

Table 5.2 – Time complexity and number of operations for one iteration of ELBND algorithms, n is the number of adaptive model parameters. The table is from [3]

The final measured results from comparison of the ELBND and the other similar methods is possible to see in Tab. 5.3.

n	ELBND	LE	FD	MD
3	0.040565	12.106009	17.114869	27.900899
23	0.067401	12.566594	132.167673	51.513525
43	0.091092	13.055424	211.303988	71.177681
63	0.112905	13.724406	291.961026	95.538772
83	0.140568	14.402364	370.774985	128.212707

Table 5.3 – Measured time for all algorithms in miliseconds.

As the results shows, the ELBND is much faster than the other methods. Especially if the big number of adaptive weights is required for evaluation.

6 | Conclusions

In this thesis the derivation, implementation and experimental analysis of newly developed adaptive novelty detection method - ELBND is described. The method is designed to by used with any supervised adaptive algorithm that has adaptive parameters and error. The experimental analysis that is present in this work features adaptive filters as the adaptive models used together with ELBND.

As shown in previous chapters, the ELBND method with adaptive filters does not require the knowledge of future samples (whole batches) [2,3,17,27]. Thus the ELBND is suitable for online data streams processing. Therefore *the 1. goal of this thesis is accomplished*. This is the key feature of the ELBND method, because not many novelty detection methods are designed to be able work in this way.

Yet another important aspect of the ELBND method is the requirement for low computational power [1–3, 16]. The time complexity of the ELBND method is only constant. In other words, the constant time complexity is the lowest time complexity possible to have for an algorithm. *As a conclusion the ELBND overall speed is good enough to accomplish the 2. goal of this thesis*. The speed aspect of the ELBND together with its ability to work without data batches makes it perfect option for real time novelty detection applications.

The ELBND is an adaptive method and this fact should ensure some concept drift robustness by itself. The ELBND ability to perform well with distorted data was investigated more via means of experimental analysis [1, 14, 28]. Although it is difficult to measure this robustness in some fair way, the results indicate that the ELBND posses ability to deal with reasonably sized concept drift. Even when tested with adaptive models derived with assumption of zero-mean data. *Therefore the 3. goal of this thesis is also accomplished*. According to the results presented, the ELBND can be considered as an algorithm safe for operation on data with reasonable sized gradual or recurring concept drift.

List of author publications

Thesis related

The following author publications are in the scope of the thesis and results have been included or cited in the text of the thesis.

A1 M. Cejnek, P. M. Benes, and I. Bukovsky, "Another adaptive approach to novelty detection in time series."   Fourth International conference on Computer Science and Information Technology, Academy and Industry Research Collaboration Center (AIRCC), 2014, pp. 341–351, doi: 10.5121/csit.2014.4229, ISBN 78-1-921987-27-4.

A2 M. Cejnek, "Rychlé algoritmy pro adaptivní detekci novosti," in *Nové metody a postupy v oblasti přístrojové techniky, automatického řízení a informatiky 2018*.   České vysoké učení technické v Praze, 2018, pp. 80 – 90, ISBN 978-80-01-06477-1.

A3 M. Cejnek and I. Bukovsky, "Influence of type and level of noise on the performance of an adaptive novelty detector," in *2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*.   IEEE, 2017, pp. 373–377, ISBN 978-1-5386-0771-8.

A4 M. Cejnek and A. Peichl, "Rychlost adaptivní algoritmů pro detekce novosti," in *Nové metody a postupy v oblasti přístrojové techniky, automatického řízení a informatiky 2018*.   České vysoké učení technické v Praze, 2018, pp. 91 – 99, ISBN 978-80-01-06477-1.

A5 M. Cejnek, I. Bukovsky, N. Homma, and O. Liska, "Adaptive polynomial filters with individual learning rates for computationally efficient lung tumor motion prediction," in *Computational Intelligence for Multimedia Understanding, 2015 International Workshop on*.   IEEE, 2015, pp. 1–5, ISBN 9781467384582.

A6 M. Cejnek and I. Bukovsky, "Online data centering modifications for adaptive filtering with nlms algorithm," in *Neural Networks (IJCNN), 2016 International Joint Conference on*.   IEEE, 2016, pp. 1767–1771, ISBN 978-1-5090-0620-5.

A7 M. Cejnek, I. Bukovsky, and O. Vysata, "Adaptive classification of eeg for dementia diagnosis," in *Computational Intelligence for Multimedia Understanding (IWCIM), 2015 International Workshop on.* IEEE, 2015, pp. 1–5, ISBN 9781467384582.

A8 M. Cejnek and I. Bukovsky, "Concept drift robust adaptive novelty detection for data streams," *Neurocomputing*, vol. 309, pp. 46 – 53, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231218305253

A9 I. Bukovsky, M. Cejnek, J. Vrba, and N. Homma, "Study of learning entropy for onset detection of epileptic seizures in eeg time series," in *Neural Networks (IJCNN), 2016 International Joint Conference on.* IEEE, 2016, pp. 3302–3305, ISBN 978-1-5090-0620-5.

Other publications

The following list of selected author publications have not been included or cited in the text.

B1 I. Bukovsky, N. Homma, K. Ichiji, M. Cejnek, M. Slama, P. M. Benes, and J. Bila, "A fast neural network approach to predict lung tumor motion during respiration for radiation therapy applications," *BioMed research international*, vol. 2015, 2015.

B2 I. Bukovsky, C. Oswald, M. Cejnek, and P. M. Benes, "Learning entropy for novelty detection a cognitive approach for adaptive filters," in *2014 Sensor Signal Processing for Defence (SSPD).* IEEE, 2014, pp. 1–5.

B3 I. Bukovsky, N. Homma, M. Cejnek, and K. Ichiji, "Study of learning entropy for novelty detection in lung tumor motion prediction for target tracking radiation therapy," in *2014 International Joint Conference on Neural Networks (IJCNN).* IEEE, 2014, pp. 3124–3129.

B4 P. M. Benes, I. Bukovsky, M. Cejnek, and J. Kalivoda, "Neural network approach to railway stand lateral skew control," *arXiv preprint arXiv:1402.7136*, 2014.

B5 M. Cejnek and C. Oswald, "Machine vision object measurement in difficult industry environment," in *2019 Third World Conference on Smart Trends in Systems Security and Sustainablity (WorldS4).* IEEE, 2019, pp. 167–170.

Bibliography

[1] M. Cejnek and I. Bukovsky, "Concept drift robust adaptive novelty detection for data streams," *Neurocomputing*, vol. 309, pp. 46 – 53, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231218305253

[2] M. Cejnek, "Rychlé algoritmy pro adaptivní detekci novosti," in *Nové metody a postupy v oblasti přístrojové techniky, automatického řízení a informatiky 2018*.   České vysoké učení technické v Praze, 2018, pp. 80 – 90, ISBN 978-80-01-06477-1.

[3] M. Cejnek and A. Peichl, "Rychlost adaptivní algoritmů pro detekce novosti," in *Nové metody a postupy v oblasti přístrojové techniky, automatického řízení a informatiky 2018*.   České vysoké učení technické v Praze, 2018, pp. 91 – 99, ISBN 978-80-01-06477-1.

[4] V. S. A. Kumar *et al.*, "Comparison of stable nlmf and nlms algorithms for adaptive noise cancellation in ecg signal with gaussian, binary and uniform signals as inputs," *International Journal of Engineering Research and Applications*, vol. 4, no. 8, pp. 28–33, 2014.

[5] V. H. Nascimento and J. C. M. Bermudez, "Probability of divergence for the least-mean fourth algorithm," *IEEE transactions on signal processing*, vol. 54, no. 4, pp. 1376–1385, 2006.

[6] P. I. Hubscher and J. C. M. Bermudez, "An improved statistical analysis of the least mean fourth (lmf) adaptive algorithm," *IEEE transactions on Signal Processing*, vol. 51, no. 3, pp. 664–671, 2003.

[7] L. Aguayo and G. A. Barreto, "Novelty detection in time series using self-organizing neural networks: A comprehensive evaluation," *Neural Processing Letters*, pp. 1–28, 2017.

[8] V. Hautamaki, I. Karkkainen, and P. Franti, "Outlier detection using k-nearest neighbour graph," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 3.   IEEE, 2004, pp. 430–433.

[9] R. O. Duda, P. E. Hart, D. G. Stork *et al.*, *Pattern classification*. Wiley New York, 1973, vol. 2.

[10] V. Barnett and T. Lewis, "Outliers in statistical data: Wiley series in probability and statistics," 1994.

[11] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Processing*, vol. 99, pp. 215–249, 2014.

[12] C. C. Aggarwal and P. S. Yu, "Outlier detection with uncertain data," in *Proceedings of the 2008 SIAM International Conference on Data Mining*. SIAM, 2008, pp. 483–493.

[13] F. E. Grubbs, "Procedures for detecting outlying observations in samples," *Technometrics*, vol. 11, no. 1, pp. 1–21, 1969.

[14] M. Cejnek and I. Bukovsky, "Online data centering modifications for adaptive filtering with nlms algorithm," in *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016, pp. 1767–1771, ISBN 978-1-5090-0620-5.

[15] M. Cejnek, I. Bukovsky, and O. Vysata, "Adaptive classification of eeg for dementia diagnosis," in *Computational Intelligence for Multimedia Understanding (IWCIM), 2015 International Workshop on*. IEEE, 2015, pp. 1–5, ISBN 9781467384582.

[16] M. Cejnek, I. Bukovsky, N. Homma, and O. Liska, "Adaptive polynomial filters with individual learning rates for computationally efficient lung tumor motion prediction," in *Computational Intelligence for Multimedia Understanding, 2015 International Workshop on*. IEEE, 2015, pp. 1–5, ISBN 9781467384582.

[17] M. Cejnek, P. M. Benes, and I. Bukovsky, "Another adaptive approach to novelty detection in time series." Fourth International conference on Computer Science and Information Technology, Academy and Industry Research Collaboration Center (AIRCC), 2014, pp. 341–351, doi: 10.5121/csit.2014.4229, ISBN 78-1-921987-27-4.

[18] D. P. Mandic, "A generalized normalized gradient descent algorithm," *Signal Processing Letters, IEEE*, vol. 11, no. 2, pp. 115–118, 2004.

[19] B. Schölkopf and A. J. Smola, *Learning with kernels: Support vector machines, regularization, optimization, and beyond.* MIT press, 2002.

[20] C. M. Bishop, "Novelty detection and neural network validation," in *Vision, Image and Signal Processing, IEE Proceedings*-, vol. 141. IET, 1994, pp. 217–222.

[21] T. Kohonen, "Self-organizing maps, vol. 30 of springer series in information sciences," 2001.

[22] S. Singh and M. Markou, "An approach to novelty detection applied to the classification of image regions," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 16, no. 4, pp. 396–407, 2004.

[23] M. Markou and S. Singh, "Novelty detection: a review—part 1: statistical approaches," *Signal processing*, vol. 83, no. 12, pp. 2481–2497, 2003.

[24] M. Markou and S. Singh, "Novelty detection: a review—part 2: neural network based approaches," *Signal processing*, vol. 83, no. 12, pp. 2499–2521, 2003.

[25] A. H. Sayed, *Fundamentals of adaptive filtering.* John Wiley & Sons, 2003.

[26] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology. Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–2049, Jun. 2000.

[27] I. Bukovsky, M. Cejnek, J. Vrba, and N. Homma, "Study of learning entropy for onset detection of epileptic seizures in eeg time series," in *Neural Networks (IJCNN), 2016 International Joint Conference on.* IEEE, 2016, pp. 3302–3305, ISBN 978-1-5090-0620-5.

[28] M. Cejnek and I. Bukovsky, "Influence of type and level of noise on the performance of an adaptive novelty detector," in *2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC).* IEEE, 2017, pp. 373–377, ISBN 978-1-5386-0771-8.

[29] I. Bukovsky, "Learning entropy: Multiscale measure for incremental learning," *Entropy*, vol. 15, no. 10, pp. 4159–4187, 2013.

[30] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 283, no. 3, pp. R789–R797, 2002.

[31] D. Bai, T. Qiu, and X. Li, "The sample entropy and its application in eeg based epilepsy detection," *Sheng wu yi xue gong cheng xue za zhi= Journal of biomedical engineering= Shengwu yixue gongchengxue zazhi*, vol. 24, no. 1, pp. 200–205, 2007.

[32] P. McSharry and G. Clifford, "Ecgsyn-a realistic ecg waveform generator," *URL http://www. physionet. org/physiotools/ecgsyn*, 2003.

[33] P. C. Mahalanobis, "On the generalized distance in statistics." National Institute of Science of India, 1936.

[34] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[35] S. Haykin, *Neural networks: a comprehensive foundation.* Prentice Hall PTR, 1994.

[36] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[37] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

Curriculum Vitae

NAME: Matouš Cejnek
ADDRESS: Sečská 15, Prague, Czech Republic
PHONE: +420 607 266 864
E-MAIL: matous.cejnek@fs.cvut.cz

Education

2014 – 2020	**CTU in Prague, Faculty of Mechanical Engineering** Doctoral Study Programme: Mechanical Engineering, Control and Systems Engineering
2012 – 2014	**CTU in Prague, Faculty of Mechanical Engineering** Master Study Programme: Mechanical Engineering, Instrumentation and Control Engineering Master Thesis: Vytěžování informací z internetu a jejich interaktivní vizualizace
2008 – 2012	**CTU in Prague, Faculty of Mechanical Engineering** Bachelor Study Programme: Mechanical Engineering, Information and Automation Technology Bachelor Thesis: Model Kostela Stětí svatého Jana Křtitele v Dolních Chabrech
2004 – 2008	**Střední technická škola v Jihlavě** Study Programme: Mechatronik

Professional experience

2017 – present	**Assistant at CTU in Prague, Faculty of Mechanical Engineering** Courses: Python for scientific computation and control, Simulations of biological systems, Počítačem podporované studium, Automatizace pro průmyslovou praxi
2014	**KPC group s.r.o. – IT developer**
2013 – 2014	**PC Help – web application development**

33

Publications

Author or co-author of 11 papers in conference proceedings, 2 papers in journals, and 1 book chapter.

Web of Science	9 records, 18 citations (without self citations), h-index 3
Scopus	11 records, 32 citations (without self citations), h-index 3

Projects

2018 – present	Inteligentní algoritmy pro detekci a potlačování nežádoucích dynamických jevů při řízení technických procesů (SGS18/177/OHK2/3T/12) Assignment: Research and pusblishing
2018	Využití strojového vidění pro automatické odměřování pri stříhání válcovaných plechu (CZ.01.1.02/0.0/0.0/17 205/0014475) Assignment: Research and development
2017 – present	Centrum pokročilých leteckých technologií (CZ.02.1.01/0.0/0.0/16 019/0000826) Assignment: Research and development
2017	Srovnání současných adaptivních metod detekce novosti v signálech (SGS17/070/OHK2/1T/12) Assignment: Research and publishing
2017	Pracoviště automatizované kontroly výstupu vibračních kruhových zásobníků (CZ.01.1.02/0.0/0.0/16 045/0008564) Assignment: Research and development
2015 – 2017	Nekonvenční a kognitivní metody zpracování signálů dynamických systémů II (SGS15/189/OHK2/3T/12) Assignment: Research and publishing
2014	Web - ICT nástroj dlouhodobého plánování podpory IT infrastruktury (EoL / EoS), Inovační voucher pro Prahu Assignment: Datamining and API development

| 2013 – 2015 | Zajištění kredibilní funkce regulace spojené s vyhledáním provozního optima ověřené na pilotních mechanicko-tepelných zařízeních nebo jejich modelech (SGS13/179/OHK2/3T/12) |
| | Assignment: Research and publishing |

| 2012 – 2014 | Nekonvenční a kognitivní metody zpracování signálů dynamických systémů (SGS12/177/OHK2/3T/12) |
| | Assignment: Research and publishing |

Additional Information

| Languages | English |
| | spoken and written, upper-intermediate |

Awards	2nd place at the Student's Conference STC 2019, FME, CTU
	Paper: Machine vision object tracking for automatic metal sheet cutting
	3rd place at the Student's Conference STC 2017, FME, CTU
	Paper: Padasip - open source library for adaptive signal processing in language Python
	1st place at the Student's Conference STC 2015, FME, CTU
	Paper: Adaptive classification of EEG for dementia diagnosis