

Response to the Reviewers' Comments

Manuscript id **NEUCOM-D-18-00164**

Authors: **Paweł Ksieniewicz, Michał Woźniak, Bogusław Cyganek, Andrzej Kasprzak, Krzysztof Walkowiak**

Title: **Data Stream Classification using Active Learned Neural Networks**

Dear Editor and Reviewers,

We would like to thank you for all your corrections and suggestions. The paper has been carefully revised, following your recommendations.

Sincerely,

Paweł Ksieniewicz, Michał Woźniak, Bogusław Cyganek, Andrzej Kasprzak,
Krzysztof Walkowiak

REVIEWER 1

It is interesting work, but a few improvements are required. general remarks

We would like to express our thanks for your valuable opinion.

1. Author declare that in-built forgetting mechanism is employed, but there are no information/discussion how control it, i.e., forgetting speed (as for windowing - we may control the forgetting by window size setting or define appropriate decaying function for weighting)

We would like to thank you for this comment. It is a crucial how control the forgetting process. As we know it strongly depends on the decision task difficulty and the size of the classification task. In this moment we recommend to control this behavior by choosing the neural network structure experimentally, i.e., in case of MLP with one hidden layer we may control this process by choosing number of neurons there. In this paper we added section *Inbuilt forgetting mechanism of artificial neural networks* where the classifier behavior depending on the various learning parameters is presented. We see that if the number of neurons is too low, then the forgetting is very fast what can cause that the model is unstable, on the other hand is the number of the neurons will be too high, then the adaptation will be slow and the classifier will be not enough sensitive to the changes. This observation is similar the conclusions presented in *Carpenter, G., Grossberg, S., and Rosen, D. 1991b. Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. Neural Networks 4, 6, 759–771*, where authors stated that online learner tuning is a kind of a tradeoff between stability and sensitivity.

2. I think that author may also discuss possibility of classifier ensemble building using the proposed classifier.

We would like to thank you for this comment. In this work we focus on the single model training for drifted data stream, but obviously we are going to extend our research in the future, e.g., by employing the ensemble approach. As the first step, we will use the proposed classifier in the WAE (Weighted Aging Ensemble) algorithm, which is developed by our team. We added this information in the last section, where the future research directions are enumerated.

minor improvements:

1. In sec. 2 authors define variable budget, but in the pseudocode *given_budget* is used.

The name of the parameter has been fixed.

2. threshold variable should be explained in pseudocode and

The short explanation has been added.

3. function label in not explained in pseudocode.

The explanation has been added.

4. quality of Fig.2 should be improved.

The mentioned figure has been removed.

REVIEWER 2

1. The forgetting mechanism is quite popular in the field of pattern classification. Related works part should review the relevant papers.

We would like to thank you for your fruitful comments. We extend the section Introduction, where the deeper discussion about forgetting mechanisms has been included.

2. It should be explained how to determine the two parameters, threshold and budget, and the values should be mentioned.

The reviewer is right that the appropriate parameter setting is crucial, but in this moment we propose to set them on the basis of computer experiments. The parameter budget is related to several factors, as what is the data stream speed or how many objects can be labeled without significant delay. The threshold is responsible for choosing interesting examples. If its value is low, then more of the objects are pretending to be labeled, what can cause that even if a given object is "interesting" it could not be labeled because of value of budget parameter. The presented results of the computer experiments may give future users the hints how to set them in the practical tasks.

3. Algorithm 1 should be illustrated in the main text, and the authors should highlight the originality and rationale with clearer statements.

We extended the description of the algorithm. The placement of the pseudocode is controlled by `\LATEX` environment, but it is close to the section, where the reference to the algorithm is.

4. Realistic dataset should be used to evaluate the performance.

The experiments have been carried out on 2 pools. The first one consists of real data streams: (*covtypeNorm*, *elecNormNew*, *poker-lsn*), while the second one includes computer generated streams. We added extended descriptions of used datasets. The problem of data stream classifier evaluation is that the community working on data stream classification suffers from the lack of the high number of real datasets, therefore artificial ones are commonly used.

5. Figures and tables should be replaced with clearer version to show up the main results. There are too many figures and tables to understand the main findings.

We were trying to improve the presentation of the experimental results. We present only selected results, while the detailed experimental results may be found in the article repository https://github.com/w4k2/active_learning

6. 3.6 Lessons Learned part is too naive to give anything significantly new.

We would like to thank you for your remarks. We extended the discussion of the results and we hope that it is much more clear.