Locating Changepoints in Multidimensional Time Series Using Non-parametric Methods

Dmitriy Klyushin and Andrii Urazovskyi

Taras Shevchenko National University of Kyiv, prospekt Glushkova, 4D, 03680, Kyiv, Ukraine

Abstract

In many fields, from finance to healthcare to engineering, there is a growing need to monitor and analyze large and complex multivariate time series. These time series often contain critical information that can be used to improve decision-making and optimize system performance. However, these time series can also be noisy and subject to various forms of interference, making it difficult to extract meaningful insights. One important challenge is identifying the moments when the underlying process changes, also known as changepoints. Detecting these changepoints in real-time is crucial for timely intervention and improved outcomes. In this paper, we explore the use of Fisher's linear discriminant and Petunin statistics for detecting changepoints in multivariate time series. We show how this approach can be applied to computer modeling and intelligent systems to improve the accuracy and efficiency of decision-making in a wide range of fields.

Keywords [1]

Time series, changepoint, nonparametric statistics, computer modeling, intelligent systems.

1. Introduction

Automatic systems and artificial intelligence can be used to recognize changepoints in multidimensional time series, providing valuable opportunities in various fields such as medicine, engineering, economics, and cybersecurity. This can help optimize the allocation of human resources, allowing them to focus on management and critical issues that affect people's lives. Nuclear power plants serve as an important example of the responsible use of computer modeling and intelligent systems. Safety is crucial in the design, use, economics, and licensing of such energy sources. To prevent and mitigate the consequences of accidents, it is essential to ensure the integrity and operability of vital elements within nuclear power plants. Designers have historically incorporated redundant and diverse safety features into these plants to provide reliability, ensuring that the health and safety of workers and the public can be protected with a high level of confidence even in abnormal and unplanned situations.

To be practical, a method should possess several characteristics, including:

1. High precision to minimize false negative and false positive outcomes.

2. Robustness to withstand individual outlying data points that may skew the entire data series and generate false changepoints.

3. Insensitivity to underlying distributions to maximize its applicability across different domains, scenarios, and objects.

4. Low computational cost to enable real-time processing without excessive resource utilization or server overload.

5. Optimal sensitivity that is neither too high to detect insignificant changes nor too low to miss critical events, such as nuclear reactor meltdowns or medical emergencies.

This paper will introduce a novel method for detecting changepoints in multivariate time series, which is based on a metric developed in a previous study [1]. This new method has been shown to

outperform the Kolmogorov-Smirnov and Wilcoxon statistics, as demonstrated in a recent study [2]. The paper will also discuss the potential applications of this method in the field of computer modeling and intelligent systems, specifically in the area of medicine. Section 2.1 will describe the algorithm used to calculate the Petunin statistic and its properties, while section 2.2 will discuss the algorithm for constructing the Fisher linear discriminant. Section 2.3 will provide an overview of the current state of research on detecting changepoints in multivariate time series. Section 3.1 will present the results of various numerical experiments involving different distributions.

Through a series of experiments and tests, we demonstrate the strength and effectiveness of our method, showing that it is capable of producing good results. Overall, we believe that our method represents a step forward in scientific research, offering a truly unique and innovative approach to solving some of the most challenging problems of our time. We hope that our work will inspire others to explore new and unconventional methods, leading to even more groundbreaking discoveries in the future.

2. Theoretical part

This chapter provides a comprehensive overview of the existing literature, as well as the statistical tools, namely Petunin's statistic and Fisher's linear discriminant analysis, that will be used in this study. This paper presents novel theoretical results that combine Fisher's linear discriminant analysis and Petunin's statistic for data analysis in our topic. This is the first study to use these two statistical tools in combination for analyzing data and our results demonstrate the unique benefits of this approach.

2.1. Petunin's statistics

Yuriy Petunin, a mathematician from Ukraine, introduced the p-statistic, which measures the closeness between two samples. The p-statistic is employed to test the hypothesis that the distribution functions of two samples are identical.

Let us consider two general populations G and G' and corresponding distribution functions F_G and $F_{G'}$.

Let there be two samples $x = (x_1, x_2, \dots x_n) \in G$ and $x' = (x_1', x_2', \dots x_m') \in G'$, and $x_{(1)}, \leq x_{(2)} \leq x_{(3)} \dots \leq x_{(n)}$ and $x_{(1)}' \leq x_{(2)}' \leq x_{(3)}' \dots \leq x_{(m)}'$ - corresponding ordinal statistics and it is necessary to determine whether they belong to the same distributions. Suppose that $F_G(u) = F_{G'}(u)$, then

$$P\left(A_{ij}^{(k)}\right) = P\left(x_k' \in \left(x_{(i)}, x_{(j)}\right)\right) = p_{ij}^{(n)} = \frac{j - i}{n + 1}$$

If we have a sample $x' \in \left(x_{(1)}', x_{(2)}', x_{(3)}', \dots, x_{(m)}'\right)$, we can find the frequency h_{ij} random event A_{ij} and confidence intervals $\left(\Delta_{ij}^{(1)}, \Delta_{ij}^{(2)}\right)$ for probability p_{ij} at a given level of significance β, i.e

$$B = \left\{p_{ij} \in \left(\Delta_{ij}^{(1)}, \Delta_{ij}^{(2)}\right)\right\}, p(B) = 1 - \beta$$

According to [4]

$$\Delta_{ij}^{(1)} = \frac{h_{ij}^{(n)} n + \frac{g^2}{2} - g\sqrt{h_{ij}^{(n)}\left(1 - h_{ij}^{(n)}\right)n + \frac{g^2}{4}}}{n + g^2}$$

$$\Delta_{ij}^{(2)} = \frac{h_{ij}^{(n)} n + \frac{g^2}{2} + g\sqrt{h_{ij}^{(n)}\left(1 - h_{ij}^{(n)}\right)n + \frac{g^2}{4}}}{n + g^2}$$

Assuming $\phi(g) = 1 - \frac{\beta}{2}$ ($\phi(g)$ where $\phi(g)$ is the normal distribution density, we can determine the significance level of the confidence interval $I_{ij}^{(n,m)} = (\Delta_{ij}^{(1)}, \Delta_{ij}^{(2)})$, using the value of g. As per the 3σ rule [5], at $g = 3$, the significance level of this interval is no more than 0.05. Let N be the total number of confidence intervals $I_{ij} = (\Delta_{ij}^{(1)}, \Delta_{ij}^{(2)})$, where $N = \frac{n(n-1)}{2}$. We define L as the number of

intervals I_{ij} that contain the probability $p_{ij}^{(n)}$. The p-statistics, $h^{(n)} = \frac{L}{N}$, is a measure of closeness $\rho(x, x')$ between samples x and x'. By substituting the obtained value of h into the formula for calculating confidence intervals, we obtain the confidence interval $I = (\Delta^{(1)}, \Delta^{(2)})$ to test the hypothesis H with an approximate significance level of 0.05 [1].

2.2. Fisher's linear discriminant

Fisher's linear discriminant and LDA (Linear discriminant analysis) are terms that are often used interchangeably, but Fisher's original paper [3] describes a discriminant that differs slightly from LDA. Fisher's method does not rely on some of the assumptions of LDA, such as classes with normal distributions or equal class covariances.

Consider two classes of observations with means $\overrightarrow{\mu_0}, \overrightarrow{\mu_1}$ and covariances Σ_0, Σ_1. If we use the linear combination of features $\vec{w} \cdot \vec{x}$, the means of the resulting distribution will be $\vec{w} \cdot \overrightarrow{\mu_i}$, and the variances will be $\vec{w}^T \Sigma_i \vec{w}$ for $i = 0,1$. Fisher defined the separation between these two distributions to be the ratio of the variance between the classes to the variance within the classes:

$$S = \frac{\sigma_{between}^2}{\sigma_{within}^2} = \frac{(\vec{w} \cdot \overrightarrow{\mu_i} - \vec{w} \cdot \overrightarrow{\mu_0})^2}{\vec{w}^T \Sigma_1 \vec{w} + \vec{w}^T \Sigma_0 \vec{w}} = \frac{\left(\vec{w} \cdot (\overrightarrow{\mu_i} - \overrightarrow{\mu_0})\right)^2}{\vec{w}^T (\Sigma_0 + \Sigma_1) \vec{w}}$$

The measure described here is a way to evaluate the effectiveness of class labelling by comparing the separation between two sets of observations to the variance within each set. The maximum separation is achieved when a certain linear combination of features is used, with the vector representing the normal to the discriminant hyperplane.

$$\vec{w} \propto (\Sigma_0 + \Sigma_1)^{-1}(\overrightarrow{\mu_i} - \overrightarrow{\mu_0})$$

In a two-dimensional problem, this hyperplane is represented by a line that is perpendicular to this vector. The data points are then projected onto this hyperplane, and a threshold is chosen based on analysis of the one-dimensional distribution of the projections. One possible way to set this threshold is by placing it between the projections of the means of the two sets of observations.

$$c = \vec{w} \cdot \frac{1}{2}(\overrightarrow{\mu_0} + \overrightarrow{\mu_1}) = \frac{1}{2}\overrightarrow{\mu_1}^T \Sigma_1^{-1}\overrightarrow{\mu_1} - \frac{1}{2}\overrightarrow{\mu_0}^T \Sigma_0^{-1}\overrightarrow{\mu_0}$$

The value of the parameter c in the threshold condition $\vec{w} \cdot \vec{x} > c$ can be explicitly determined in this case. It should be noted that Fisher's original discriminant differs slightly from LDA in terms of the assumptions it makes about the classes being normally distributed or having equal covariances.

2.3. Related works

There are numerous approaches to detection of changepoints in multidimensional time series of random values. A changepoint of a time series is such point where values of time series have different distribution that before and after that point. Methods for detection changepoints in multidimensional time series can be classified into online and offline algorithms. Online algorithms work on portion of data in time, but offline algorithms work on completed sets of data. Article [6] provided a comprehensive review of offline methods for changepoint detection. For online algorithms, we aim to consider a method that is independent of initial distributions.

The article [7] proposes a novel approach for discriminant analysis, called Kernel Fisher Discriminant, which shows competitive performance compared to other classification techniques and has potential for further extensions in multi-class discriminants and generalization error bounds.

Increasing of dimension can slow down computations. Article [8] discussed application of divergence measures to detect a changepoint in a time series. Changepoint detection can be performed in different ways, such as detection of a changepoint in time series or localization of its coordinates. Article [9] focused on detecting a changepoint, but their method has not high accuracy in localization. Article [10] proposed a Bayesian method with linear computational complexity, but its accuracy is insufficient. Article [11] developed an effective convex network clustering algorithm, but it is computationally complex. The article [12] proposes a change-point based control chart for monitoring sparse changes in high-dimensional mean vector in HDLSS scenarios, which is robust to correlation,

non-normality, and heteroscedasticity and shows efficient detection of large sparse shifts with accurate estimation of the change-point and potential OC variables, as shown by experimentation and a real case study.

To process online data, identify outliers without restricted assumptions about the data distribution, we will examine papers that consider our problem from the same point of view. Articles [13] and [14] developed a Bayesian method for exploring geographical data. Article [15] considered algorithms for exponential models only, while article [16] requires information about type of distribution to increase the accuracy of their method. Article [17] made assumptions on distribution to decrease computational complexity. Article [18] also required prior assumptions about the data. Pre-processing the data can increase the precision of changepoint detection in multivariate time series [19]. Articles [20] and [21] proposed Bayesian methods for segmenting multivariate time series with implicit examination of a dependency structure.

In the study by [22], an algorithm was examined for streaming data, which relied on a massive matrix that was contingent on the size of the original data space. Comparable techniques were explored in a different research paper by Romano and others in [23].

The Petunin's statistics is a measure of distance between two distributions, which can be used to detect a changepoint in a time series. The proposed algorithm is based on this statistic and has the following properties:

1. Stability: The algorithm is designed to be stable over time, meaning that it can accurately detect changepoints even when underlying distribution of data changes over time.

2. High accuracy: The Petunin's statistics is a robust measure of distance between distributions. It allows accurate detection of changepoints even in the presence of outliers or other noise in data.

3. Speed: The algorithm is designed to be computationally efficient, allowing for real-time processing of streaming data.

4. Independence from basic distributions: The algorithm does not require any assumptions or prior knowledge about the underlying distribution of the data, making it applicable to a wide range of time series data.

In summary, the proposed algorithm based on Petunin's statistic offers a stable, accurate, and computationally efficient method for detecting changepoints in time series data, without requiring any specific assumptions about the underlying distribution of data.

3. Practice part

In this chapter, we present a method for detecting changepoints in time series data and conduct several numerical experiments to evaluate its performance with different distributions.

Our method is based on a combination of statistical tools, including Petunin's statistic and Fisher's linear discriminant analysis. By using these tools in combination, we can identify changepoints in time series data with high accuracy.

The purpose of our experiments is to demonstrate the accuracy of the following algorithm for a stationary time series, which should find the first changepoint and test the homogeneity hypothesis.

At the beginning we take $width$ and designate the elements $x_1, \ldots, x_{width}$ - starting ones, with which we will continue to work using the sliding window method. When we have a sample $(x_{i+1}, x_{i+2}, \ldots, x_{i+width})$, we do the following with it:

- Building a linear Fisher discriminant for samples $(x_1, x_2, \ldots, x_{width})$ and $(x_{i+1}, x_{i+2}, \ldots, x_{i+width})$ and find the projections on the line.
- Rotate the resulting straight line so that only one coordinate remains, and make the rest the same. Getting projections $(p_1, p_2, \ldots, p_{width})$ and $(p_{i+1}, p_{i+2}, \ldots, p_{i+width})$
- Calculate the Petunin's statistics p_{stat} for the resulting sets of projections
- If $p_{stat} \geq 0.95$, then we say that the new sample has the same distribution as the original one, otherwise we say that the other and shift the initial sample to position $(p_{i+width+1}, \ldots, p_{i+2 \cdot width})$.
- Shifting our window one position to the right and start the algorithm from the beginning. We do this until all the data is gone.

If sample after element x_n become inhomogeneous, then the point x_{n+1} regarded as a changepoint.

To demonstrate how the algorithm works, we take a series of length $N = 400$ and divide it into 4 equal intervals with different distributions. Then we run our algorithm 100 times and average the values of Petunin's statistics (P statistics), after which we display the obtained values in two colors: blue is not less than 0.95, that is, for those samples that have the same distribution as the original and red less than 0.95 - having a different distribution.

For each experiment, we calculated five measures of error: mean absolute error (MAE), mean squared error (MSE), mean squared deviation (MSD), root mean squared error (RMSE), and normalized root mean squared error (NRMSE). To demonstrate the effectiveness of the described algorithm, we will rely on the latter value. As is well known, if NRSME > 0.5 the results can be considered as random. If a NRMSE is close to 0, then the results are considered good.

We aim to conduct a numerical experiment that involves analyzing several time series with jumps of varying distributions. In the first scenario, we will analyze a time series consisting of nearly non-overlapping uniform distributions and test the hypothesis of a shift in this series. In the second scenario, we will analyze a time series with jumps of uniform distributions initially exhibiting significant overlap, followed by mild overlap, and ultimately no overlap, with the purpose of testing the hypothesis of a shift. Additionally, we will examine time series with jumps of normal distributions, where distinct means exhibit minimal overlap, to test the hypothesis of a shift. In the fourth scenario, we will examine time series with jumps that comprise normal distributions with identical means but gradually differing variances, aiming to test the scale hypothesis. Finally, in the fifth scenario, we will analyze a time series with jumps of normal distributions with the same means but differing variances, testing the scale hypothesis on this series.

3.1. Nearly non-overlapping uniform distributions with different means

We will analyze time series with jumps that consists of nearly non-overlapping uniform distributions. The aim is to test the hypothesis of a shift in this series.

Table 1
Time intervals and uniform distributions with different means

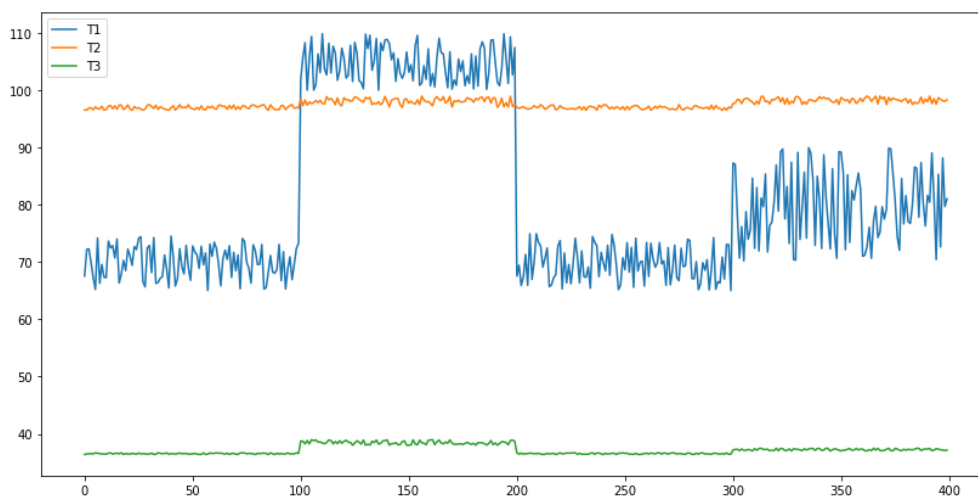| Time interval	Distribution T_1	Distribution T_2	Distribution T_3
0-99	U(65;75)	U(96.5;97.5)	U(36.4;36.7)
100-199	U(100;110)	U(97.0;99.0)	U(38.0;39.0)
200-299	U(65;75)	U(96.5, 97.5)	U(36.4;36.7)
300-399	U(70,90)	U(97.5;99.0)	U(37.0;37.5)

Figure 1: Time series composed of samples from nearly non-overlapping uniform distributions with varying means and their respective changepoints.
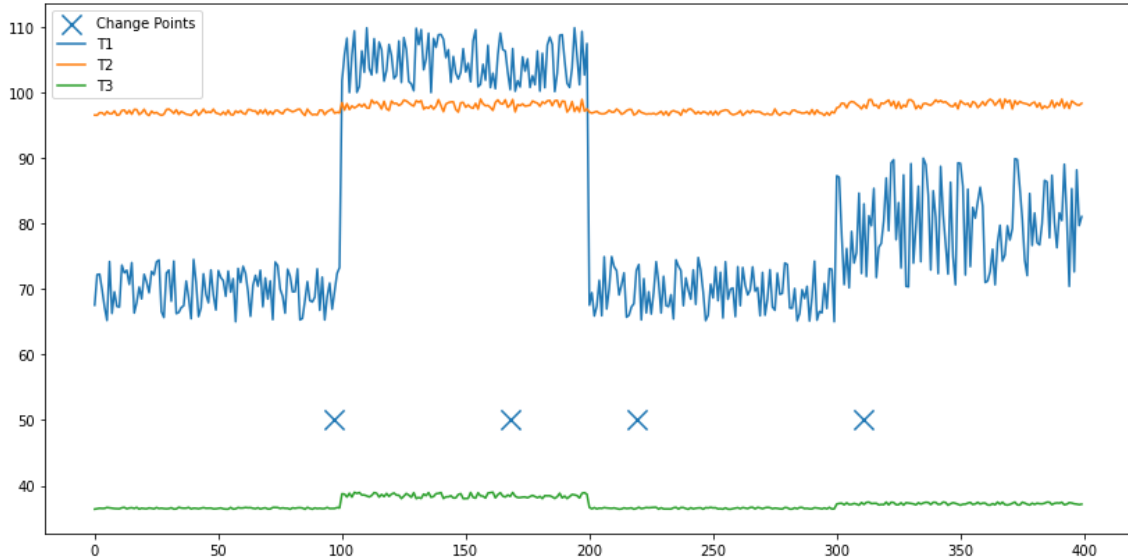
Figure 2: Time series composed of samples derived from nearly non-overlapping uniform distributions that have distinct means and corresponding changepoints, as denoted by blue crosses using the algorithm.

Table 2

Measures of error for nearly non-overlapping uniform distributions that have distinct means.

Error measure	Value
MAE	44.74
MSE	2002.73
MSD	20.73
RMSE	42.41
NRMSE	0.21

Table 1 and Figure 1 illustrate that the intended changepoints are 100, 200 and 300. In Figure 2, we saw that the almost all found changepoints are close to the actual ones, while the measures of error are presented in Table 2.

3.2. Uniform distributions with distinct means that display significant overlap at the outset, followed by mild overlap, and ultimately no overlap

We will analyze time series with jumps that consists of uniform distributions initially exhibiting significant overlap, followed by mild overlap, and ultimately no overlap. The purpose is to test the hypothesis of a shift in this series.

Table 3

Time intervals and uniform distributions with distinct means that initially display significant overlap, followed by mild overlap, and ultimately no overlap

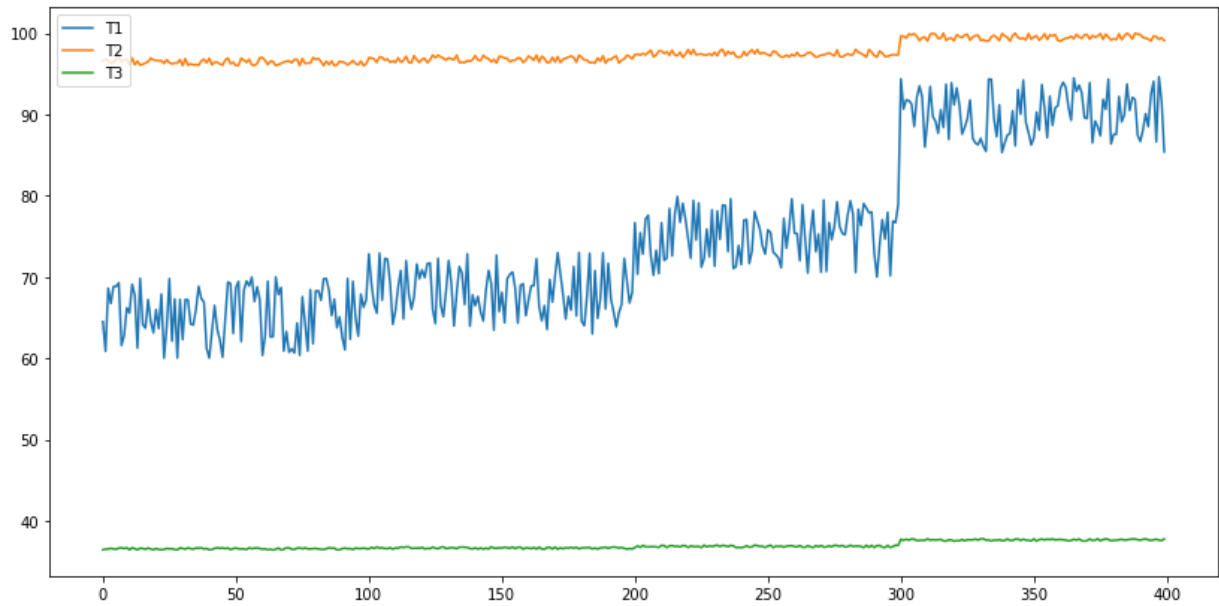| Time interval	Distribution T_1	Distribution T_2	Distribution T_3
0-99	U(60;70)	U(96.0;97.0)	U(36.4;36.7)
100-199	U(63;73)	U(96.3;97.3)	U(36.5;36.8)
200-299	U(70;80)	U(97.0, 98.0)	U(36.7;37.0)
300-399	U(85,95)	U(99.0;99.9)	U(37.5;37.8)

Figure 3: Time series composed of samples derived from uniform distributions that exhibit distinct means, initially showing significant overlap, followed by mild overlap, and ultimately no overlap with respective changepoints.
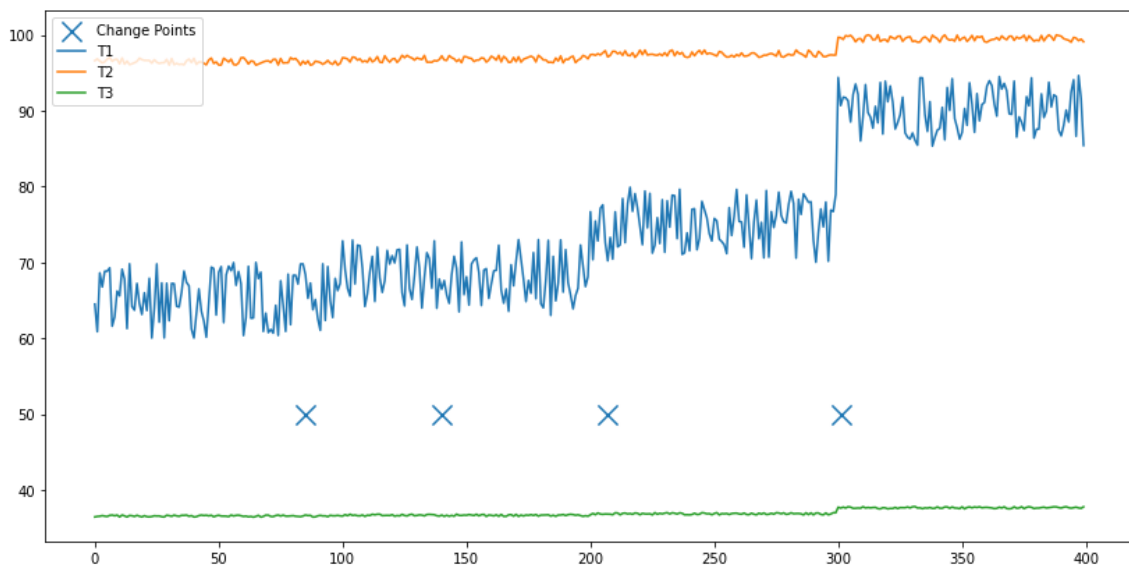

Figure 4: Time series composed of samples derived from uniform distributions that exhibit distinct means, initially showing significant overlap, followed by mild overlap, and ultimately no overlap and corresponding changepoints, as denoted by blue crosses using the algorithm.

Table 4

Measures of error for uniform distributions with distinct means that exhibit significant overlap initially, followed by mild overlap, and ultimately no overlap.

Error measure	Value
MAE	44.68
MSE	2071.37
MSD	20.67
RMSE	43.24
NRMSE	0.21

Figure 3 and Table 3 show that the desired changepoints are 100, 200, and 300. Figure 4 demonstrates that the almost all detected changepoints are close to the actual ones, and the corresponding error measures are displayed in Table 4.

3.3.   Normal distributions with distinct means that exhibit minimal overlap

We will analyze time series with jumps composed of normal distributions with distinct means that exhibit minimal overlap. The aim is to test the hypothesis of a shift in this series.

Table 5
Time intervals and normal distributions with distinct means that exhibit minimal overlap.

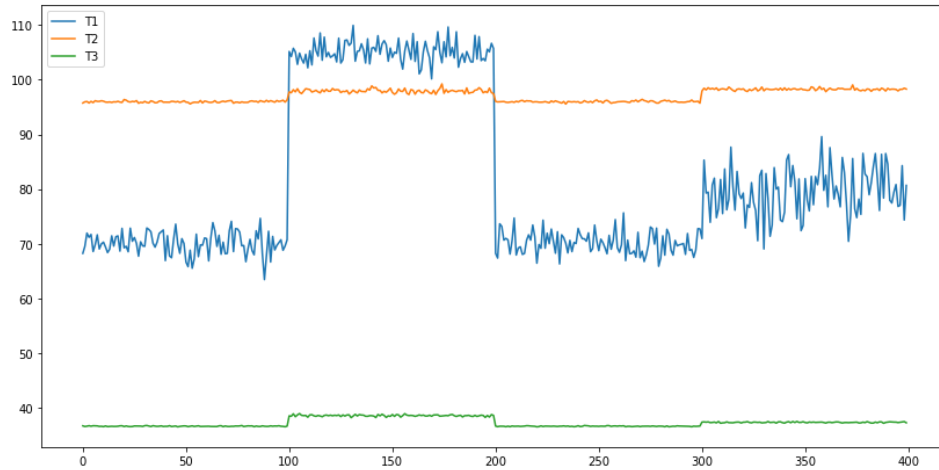| Time interval	Distribution T_1	Distribution T_2	Distribution T_3
0-99	N(70;2)	N(96.0;0.15)	N(36.5;0.05)
100-199	N(105;2)	N(96.3;0.33)	N(38.5;0.15)
200-299	N(70;2)	N(97.0, 0.15)	N(36.5;0.05)
300-399	N(80,4)	N(99.0;0.25)	N(37.3;0.98)

Figure 5: Time series composed of samples derived from normal distributions that exhibit distinct means and almost no overlap.
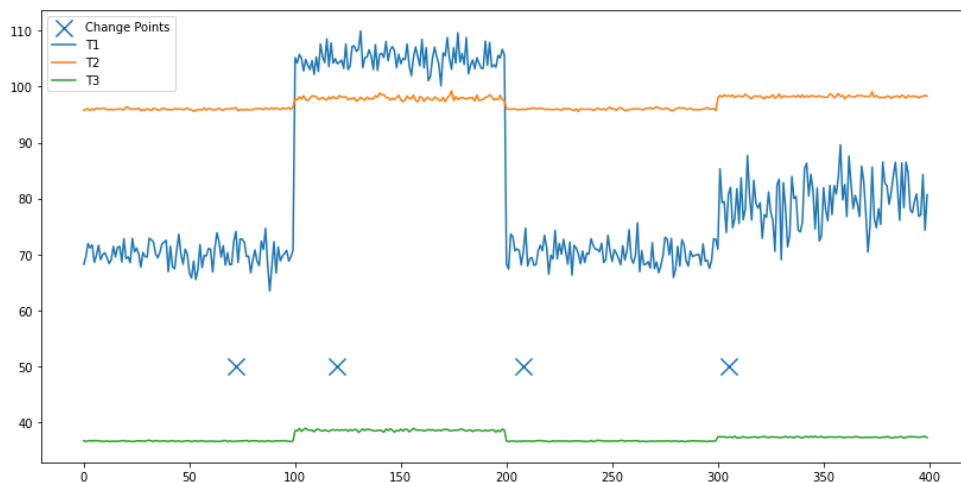

Figure 6: Time series composed of samples derived from normal distributions that exhibit distinct means and almost no overlap and corresponding changepoints, as denoted by blue crosses using the algorithm

Measures of error for normal distributions with distinct means that exhibit minimal overlap.

Error measure	Value
MAE	43.67
MSE	2077.24
MSD	22.14
RMSE	42.81
NRMSE	0.21

Table 5 and Figure 5 indicate that the intended changepoints are 100, 200, and 300. In Figure 6, we observe that the almost all detected changepoints are in close proximity to the true ones, and Table 6 displays the corresponding error measures.

3.4. Normal distributions with identical means, but whose variances begin to differ gradually

We will examine time series with jumps that comprises normal distributions with identical means, but with gradually differing variances. The objective is to test the scale hypothesis on this series.

Table 7
Time intervals and normal distributions with same means, but whose variances gradually begin to differ.

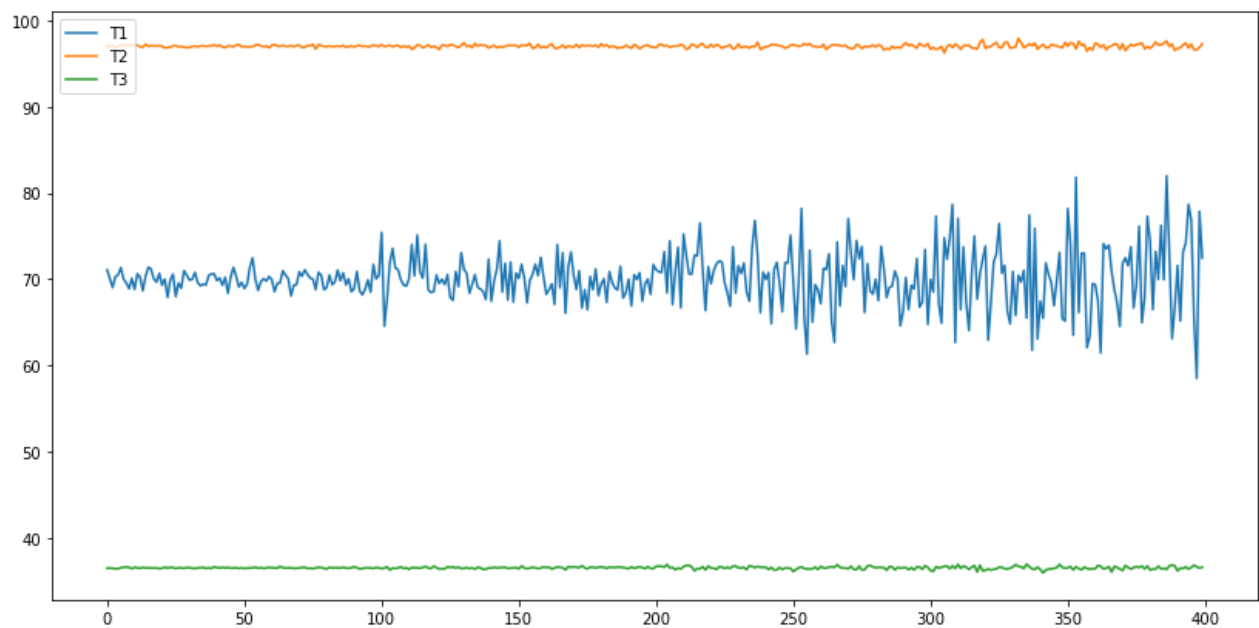| Time interval	Distribution T_1	Distribution T_2	Distribution T_3
0-99	N(70;1)	N(97.0;0.10)	N(36.55;0.05)
100-199	N(70;2)	N(97.0;0.15)	N(36.55;0.10)
200-299	N(70;3)	N(97.0,0.20)	N(36.55;0.15)
300-399	N(70;5)	N(97.0,0.30)	N(36.55;0.20)

Figure 7: Time series consisting of samples from normal distributions with the same means, but with variances that gradually begin to differ
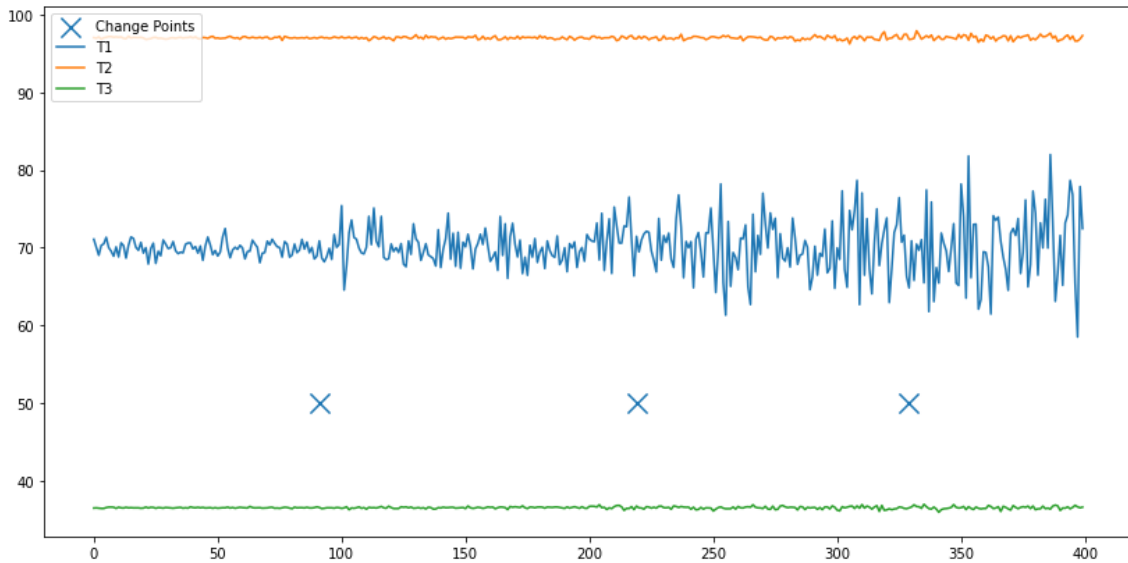
Figure 8: Time series composed of samples derived from normal distributions with identical means, but whose variances begin to differ gradually and corresponding changepoints, as denoted by blue crosses using the algorithm

Table 8

Error measures for normal distributions with the same means, but with variances that gradually begin to differ

Error measure	Value
MAE	47.75
MSE	2468.69
MSD	19.96
RMSE	46.73
NRMSE	0.23

Table 7 and Figure 7 indicate that the intended changepoints are 100, 200, and 300. In Figure 8, we observe that the almost all detected changepoints are in close proximity to the true ones, and Table 8 displays the corresponding error measures.

3.5. Normal distributions with the same means, but with variances that differ more strongly

Let's consider time series with jumps, which is composed of normal distributions with the same means, but with variances that differ more strongly. On this time series, we will be able to test the scale hypothesis.

Table 9

Time intervals and normal distributions with the same means, but with variances that differ more strongly

Time interval	Distribution T_1	Distribution T_2	Distribution T_3
0-99	N(70;1)	N(97.0;0.10)	N(36.55;0.05)
100-199	N(70;5)	N(97.0;0.50)	N(36.55;0.25)
200-299	N(70;7)	N(97.0,1.00)	N(36.55;0.5)
300-399	N(70;10)	N(97.0,1.50)	N(36.55;0.75)

As can be seen from Table 9 and Figure 9, the desired changepoint is 100. In Figure 10, we see that the p-statistic takes values greater than 0.95 only in the first interval and the measures of error we can see in Table 10.
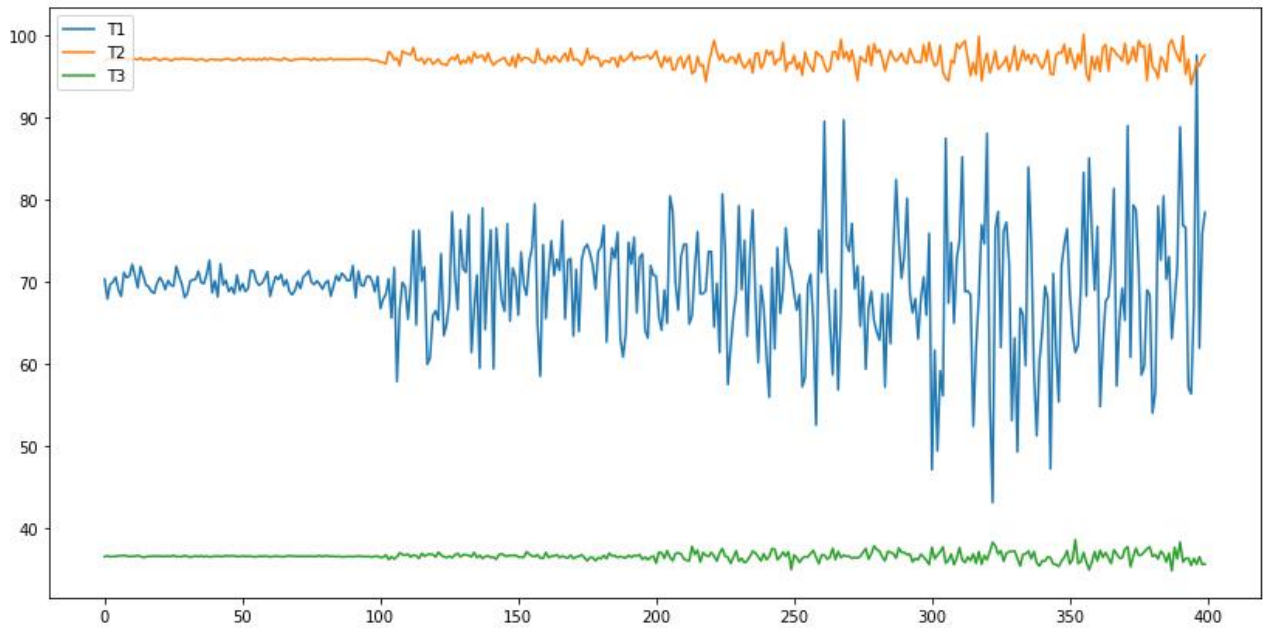

Figure 9: Time series consisting of samples from normal distributions with the same means, but with variances that differ more strongly
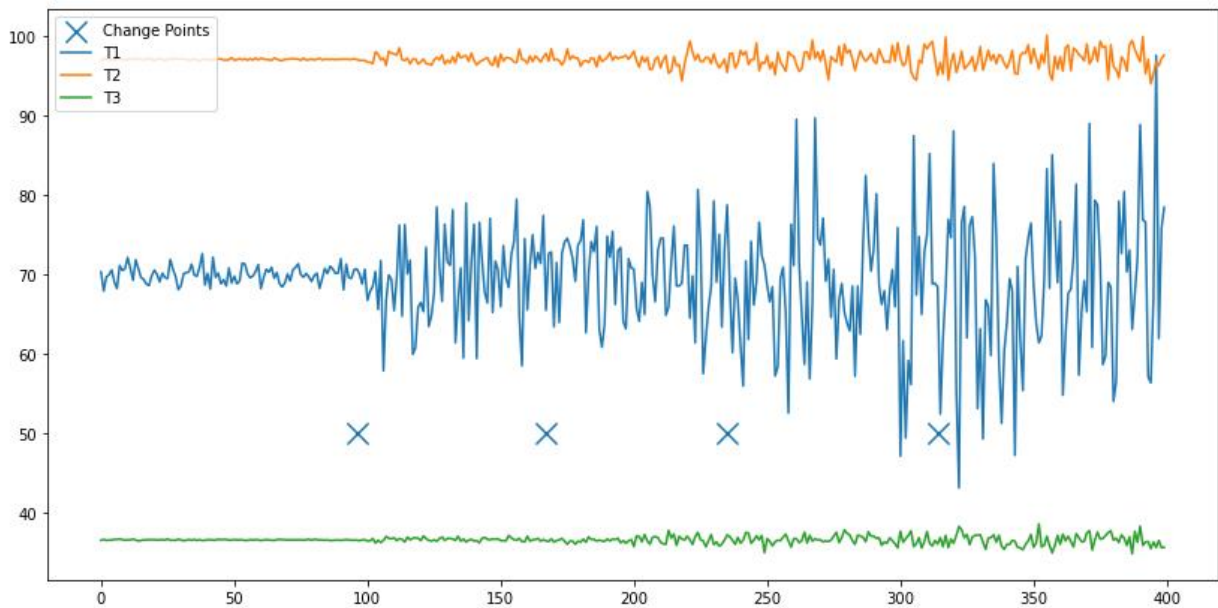

Figure 10: Time series composed of samples derived from normal distributions with the same means, but with variances that differ more strongly and corresponding changepoints, as denoted by blue crosses using the algorithm

Indeed, it is worth noting that hypotheses about scale are generally more challenging to test than those about shift. However, our algorithm is designed to detect changepoints even in scenarios where the scale hypothesis is being tested, allowing for a comprehensive analysis of the time series. By identifying these points, we can gain valuable insights into the behavior of the series and validate or reject our hypotheses.

Table 10

Error measures for normal distributions with the same means, but with variances that differ more strongly

Error measure	Value
MAE	43.49
MSE	2166.74
MSD	20.34
RMSE	43.19
NRMSE	0.22

Table 9 and Figure 9 indicate that the intended changepoints are 100, 200, and 300. In Figure 10, we observe that the almost all detected changepoints are in close proximity to the true ones, and Table 10 displays the corresponding error measures.

4. Conclusion

In conclusion, our study has presented a novel algorithm for detecting changepoints in time series data that combines Fisher's linear discriminant and Petunin's statistics. Our numerical experiments have shown that this algorithm can accurately and quickly detect changes in a wide range of distribution functions.

Additionally, our algorithm has several advantages over existing changepoint detection methods. Firstly, our algorithm does not require any assumptions about the distribution of the data, making it more flexible and applicable to a wider range of scenarios. Secondly, the computational complexity of our algorithm is relatively low, which makes it efficient and scalable to larger datasets. Finally, our algorithm provides interpretable results, which can help researchers and practitioners to better understand the nature of changes in the time series data.

The implications of our results are significant, as our algorithm could have practical applications in monitoring the health status of COVID-19 patients in clinics. By accurately detecting changes in vital signs or symptoms, medical professionals could intervene earlier and improve patient outcomes. Furthermore, we have evaluated the performance of our algorithm using NRMSE, which measures the accuracy of the detected changepoints. Our NRMSE values demonstrate that our algorithm works accurately.

However, we acknowledge that there are limitations to our study, such as using simulated data in our experiments. Therefore, the performance of our algorithm may differ when applied to real-world data. Nevertheless, our algorithm provides a valuable contribution to the field of changepoint detection, and we plan to evaluate its performance on real-world data in future research.

We hope that our combination of Fisher's linear discriminant and Petunin's statistics will inspire further research in this area and contribute to improving the accuracy and efficiency of changepoint detection algorithms. Overall, our study provides a promising foundation for future research in this field.

5. References

[1] D.A. Klyushin, Y.I. Petunin, Nonparametric population equivalence test based on measure of closeness between samples, Ukrainian Mathematical Journal (2003), 2nd. ed., pp. 147-163.
[2] D.A. Klyushin, A.V. Urazovskyi, Nonparametric Test for Change-Point Detection of IoT Time-Series Data, Chapter in: Kumar P., Obaid A., Cengiz K., Balas A. (Eds.) A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems, Intelligent Systems Reference Library, volume 210, Springer, 2021, pp. 99-122.

[3] R. A. Fisher, The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics, volume 7, 2nd. ed., 1936, pp. 179–188.

[4] B.L. Van der Waerden, Mathematische Statistic, Springer-Verlag, Berlin, 1957; English. transl. of 2nd. ed. (1965) Springer-Verlag, Berlin and New York, 1969

[5] Y. I. Petunin, D. A. Klyushin, K. P. Ganina, N. V. Borodai, R. I. Andrushkiv, Computer diagnosis of breast cancer, Bulletin of Kyiv University, Ser. cybernetics, volume 2, 2001, pp. 58-68.

[6] C. Truong, L. Oudre, N. Vayatis, Selective review of offline changepoint detection methods, Signal Processing, volume 167, 2020, 107299. doi:10.1016/j.sigpro.2019.107299.

[7] S. Mika, Fisher Discriminant Analysis with Kernels, IEEE Conference on Neural Networks for Signal Processing IX, 1999, pp. 41–48. doi:10.1109/NNSP.1999.788121

[8] C. Alippi, G. Boracchi, D. Carrera M. Roveri, Change Detection in Multivariate Datastreams: Likelihood and Detectability Loss, Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16), 2016, pp. 1368–1374. doi:10.48550/arXiv.1510.04850.

[9] Z. Wang, X. Lin, A. Mishra, R. Sriharsha, Online Changepoint Detection on a Budget, 2021 International Conference on Data Mining Workshops (ICDMW), 2021, pp. 414–420. doi:10.1109/ICDMW53433.2021.00057.

[10] S. Jaehyeok, A. Ramdas, and A. Rinaldo, E-detectors: a nonparametric framework for online changepoint detection, arXiv preprint arXiv:2203.03532v1, 2022. doi:10.48550/arXiv.2203.03532.

[11] O. Sorba, C. Geissler, Online Bayesian inference for multiple changepoints and risk assessment. arXiv preprint arXiv:2106.05834v1, 2021. doi:10.48550/arXiv.2106.05834.

[12] M. Navarro, G. I. Allen, M. Weylandt, Network Clustering for Latent State and Changepoint Detection, 2021. arXiv preprint arXiv:2111.01273v1. doi:10.48550/arXiv.2111.01273.

[13] Z. Wang, I. M. Zwetsloot, A Change-Point Based Control Chart for Detecting Sparse Changes in High-Dimensional Heteroscedastic Data, 2021. arXiv preprint arXiv:2101.09424v1. doi:10.48550/arXiv.2101.09424.

[14] L. Wendelberger, J. Gray, B. Reich, A. Wilson, Monitoring Deforestation Using Multivariate Bayesian Online Changepoint Detection with Outliers, 2021. arXiv preprint arXiv:2112.12899v2.

[15] P. Adams, D. Mackay, Bayesian Online Changepoint Detection, 2007. arXiv preprint arXiv:0710.3742v1. doi:10.48550/arXiv.0710.3742

[16] P. Cooney, A. White, Change-point Detection for Piecewise Exponential Models, 2021. arXiv preprint arXiv:2112.03962v1. doi:10.48550/arXiv.2112.03962

[17] J. Castillo-Mateo, Distribution-Free Changepoint Detection Tests Based on the Breaking of Records, 2021. arXiv preprint arXiv:2105.08186v1. doi:10.48550/arXiv.2105.08186.

[18] K. L. Hallgren, N. A. Heard, M. J. M. Turcotte, Changepoint detection on a graph of time series, 2021. arXiv preprint arXiv:2102.04112v1. doi:10.48550/arXiv.2102.04112

[19] A. Fotoohinasab, T. Hocking, F. Afghah, A Greedy Graph Search Algorithm Based on Changepoint Analysis for Automatic QRS Complex Detection, Computers in Biology and Medicine, 2021, volume 130, 104208. doi:10.1016/j.compbiomed.2021.104208

[20] P. Fearnhead, G. Rigaill, Changepoint Detection in the Presence of Outliers, Journal of the American Statistical Association, 2018, volume 114, pp. 169–183. doi:10.1080/01621459.2017.1385466

[21] F. Harlé, F. Chatelain, C. Gouy-Pailler, S. Achard, Rank-based multiple change-point detection in multivariate time series, 22nd European Signal Processing Conference (EUSIPCO), 2014, pp. 1337–1341. doi:10.5281/zenodo.43927.

[22] K. Renz, N. C. Stache, N. Fox, G. Varol, S. Albanie, Sign Segmentation with Changepoint-Modulated Pseudo-Labelling, 2021. arXiv preprint arXiv:2104.13817v1. doi:10.48550/arXiv.2104.13817

[23] G. Romano, I. Eckley, P. Fearnhead, G. Rigaill, Fast Online Changepoint Detection via Functional Pruning CUSUM statistics, 2021. arXiv preprint arXiv:2110.08205v2. doi:10.48550/arXiv.2110.08205.