



# Online Classification of High Gamma Dose Rate Incidents <sup>†</sup>

Mohammed Al Saleh <sup>1,2,3,\*</sup> , Beatrice Finance <sup>1</sup>, Yehia Taher <sup>1</sup>, Ali Jaber <sup>2</sup>  and Roger Luff <sup>4</sup>

<sup>1</sup> David Laboratory, University of Paris-Saclay, 45 Avenue des Etats-Unis, 78035 Versailles, France; beatrice.finance@uvsq.fr (B.F.); yehia.taher@uvsq.fr (Y.T.)

<sup>2</sup> Rafic Hariri University Campus, Lebanese University, Beirut 6573, Lebanon; ali.jaber@ul.edu.lb

<sup>3</sup> Lebanese Atomic Energy Commission (LAEC), National Council for Scientific Research (CNRS), Airport Road, Beirut 2260, Lebanon

<sup>4</sup> Federal Office for Radiation Protection (BfS), 24768 Rendsburg, Germany; rluff@bfs.de

\* Correspondence: mhdalsaleh@gmail.com

<sup>†</sup> Presented at the 8th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 27–30 June 2022.

**Abstract:** In this paper, we propose a new method for choosing the most suitable time-series classification method that can be applied to online gamma dose rate incidents. We referred to the historical incidents measured in the German Radiation Early Warning Network and clustered them into several classes before testing existing classification methods. This raises the research problem of the online classification of time-series data with varying scales and lengths. Referring to the state-of-the-art methods, we found that no specific classification method can fit our data all the time. This motivated us to introduce our own approach.

**Keywords:** machine learning algorithms; predictive model; time-series classification; gamma dose rate; Radiation Early Warning Network



**Citation:** Al Saleh, M.; Finance, B.; Taher, Y.; Jaber, A.; Luff, R. Online Classification of High Gamma Dose Rate Incidents. *Eng. Proc.* **2022**, *18*, 28. <https://doi.org/10.3390/engproc2022018028>

Academic Editors: Ignacio Rojas, Hector Pomares, Olga Valenzuela, Fernando Rojas and Luis Javier Herrera

Published: 22 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Time-series analysis is gaining more and more interest in so many domains. That is because, with the proliferation of the use of sensors and IoT devices that continuously produce massive amounts of real-time data, special care has been given for analyzing the data to understand past events and patterns and predict future ones. Medical heart monitor data, stock market prices, weather conditions, etc., are all examples of such time-series data.

In this paper, we are interested in analyzing the gamma dose rate (background radiation level) in the environment. A serious event that occurs and causes an abrupt increase in the gamma dose rate is the leakage or contamination failure of a nuclear reactor, such as what happened in the Chernobyl accident which was the biggest short-term leak of radioactive materials ever recorded in history [1]. Such an event has to be intercepted at the earliest point possible to take the proper measures and precautions and notify the concerned authorities to minimize the effects of such a hazardous situation. It is a very critical task as long-term or acute exposure to a high gamma dose rate can have many hazardous consequences on humans as well as on the ecosystem.

Around the globe, there are thousands of probes (sensors) that collect gamma dose rates in real time. A Radiation Early Warning System (REWS) [2] collects the data and raises the alarm in case of an increase in the local gamma dose rate. Whenever an event occurs (i.e., the gamma dose rate goes above the accepted threshold, provided by experts), an alarm is triggered, and a team of experts and personnel have to unite to investigate the reasons behind this rise. Currently, the analysis of incoming incidents is performed manually with the efforts of experts. Such a method is time-consuming and risky, knowing that the factors affecting the gamma dose rate are not always known immediately. Fortunately, most of the incoming incidents are mainly innocent ones as they remain in an acceptable range value for human health and this value returns to normal after a period of time.

The objective of our research is to propose an *Intelligent Radiation Early Warning System* that finds the cause automatically behind an incident and its classification into *real* or *innocent* ones at real time. Gaining intelligence is the key aspect of our approach. Therefore, we aimed to transform the static and semi-automatic REWSs into dynamic, fully automatic, and intelligence-driven systems. The proposed system would optimize the ability to analyze any alert generated by an event such as rain.

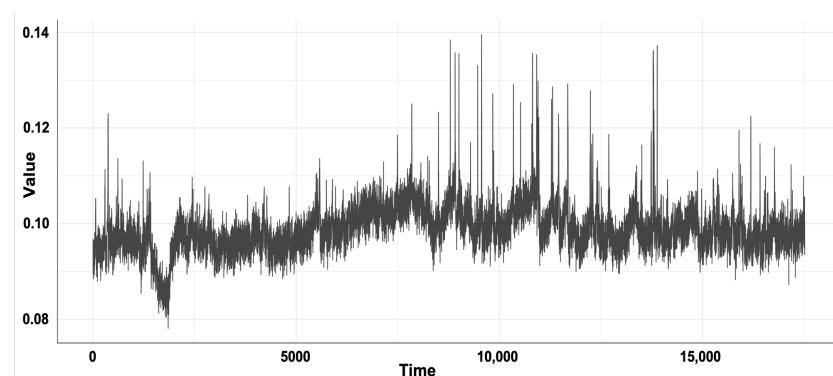
The two main phases of our *Intelligent REWS* are: (1) building the predictive model and (2) near real-time detection and prediction. In the first phase, the historical data generated by Germany's REWS [3] were analyzed to extract knowledge about the previous incidents that occurred in the past. The historical databases contain raw unlabeled data (i.e., time series) corresponding to the gamma dose rate monitoring at each probe. The data used in this work comprise the past ten years' minute-by-minute gamma dose rate real data for over a thousand probes.

In [4], we already proposed an unsupervised machine learning model that helps us automatically determine the reasons behind the incidents. This task was difficult and required many experiments to find the best time-series clustering algorithm. After tackling all the shortcomings behind the first phase of our *Intelligent REWS*, we now investigate in this paper our contributions for the Online Detection and Prediction phase. As we aim to match unlabeled incidents without any human intervention as soon as possible, our research is in the field of supervised machine learning time-series classification.

The remaining sections of this paper will be organized as follows: Section 2 will state the context and the problems behind our research. The state-of-the-art approaches, similar to our approach in one aspect or another, are described in detail in Section 3. Section 4 will present our approach and contribution. We will evaluate our approach in Section 5. Finally, we will conclude in Section 6.

## 2. Context and Problem Statement

In this research, we deal with univariate time-series which are unlabeled as shown in Figure 1. Incidents caused by the same event may not have a recognizable temporal trace or characteristics but more common behavior. For example, a particular event may cause peaks of increasing amplitudes that decrease during a longer period of time; another may cause an abrupt increase and maintain its amplitude during a period of time, and so on. Note that incidents caused by the same event can last for a varying length of time and reach different amplitudes.



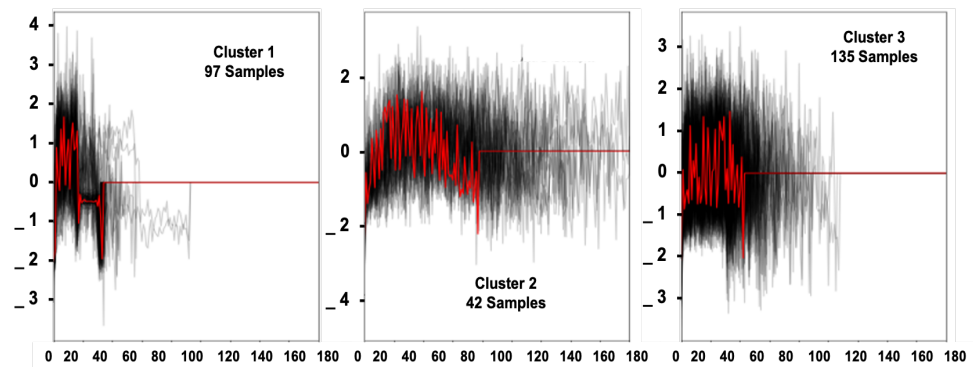
**Figure 1.** Typical gamma dose rate time series.

In the first phase, we were able to collect nearly 300 innocent incidents from 45 different locations in Germany. Choosing different locations allowed us to have diverse shapes of observations representing the innocent incidents, thus gaining a higher quality dataset to build our investigation upon. Investigating our time-series data revealed important characteristics that need to be handled carefully. Incidents are of highly varying lengths, dif-

ferent scales, and of different levels. The same incidents could have different characteristics. Going further, different incidents could have the same characteristics.

The evolving parameters problem was solved by preparing a catalog of parameters before the extraction process started. A specific algorithm tackled the evolving issue through several calculation steps to ensure that the parameters obtained by the end of each month were accurate. Then, the extracted incidents undergo a unique preprocessing phase to ensure that they are ready to enter the proposed clustering model. This was done using a *z-normalization* method [5] responsible for dealing with the scale issue. The *zero-padding* method was applied to deal with the different length incidents issue.

Once the preprocessing was applied, the best clustering model for our context was formed by combining the similarity measure (DTW) [6] as well as the clustering algorithm (K-means) [7] with its averaging method (DBA). With the help of experts in gamma radiation monitoring, we were able to identify three events after applying our clustering approach that split into three categories: rain, stormy rain, or incidents caused by probe calibration as depicted in Figure 2.



**Figure 2.** Clusters obtained from our predictive model.

For the online detection and prediction phase, a matching process tries to classify the incoming incidents as depicted in Figure 1 within the labeled clusters depicted in Figure 2. This helps identify the real cause of the current incident. The incoming readings are analyzed to explore the thresholds in real-time. Once an incident is detected, the data preprocessing model used in phase one is also applied to deal with the scale and length issues. Notice that we start analyzing an incident even if the incident is not yet finished. As we aim to match unlabeled incidents without any human intervention as quickly as possible, our research is in the field of supervised machine learning time-series classification.

After investigating the classification algorithms presented in the literature, we noticed several shortcomings that prevented us from relying on a specific algorithm for our online classification model. After going through the classification algorithms introduced in the literature, we noticed that no specific algorithm could perfectly fit our data since our data have unique characteristics and behavior. We noticed that although some algorithms work perfectly for specific types of incidents, they could not classify other types. This problem was enough for us to not trust a specific classification algorithm when dealing with our incoming incidents. Moreover, we noticed that these algorithms were unable to detect incoming incidents that have a unique behavior and should be classified in a new class that will be labeled by the experts later.

In this context, the problem consists of finding a machine-learning-based framework to automate the event identification process to decrease the time and effort spent and increase the efficiency and accuracy of the process. This will result in automatically identifying the incoming incidents as soon as possible and giving the correct impression to the experts to distinguish the innocent incidents from those that are critical. Hence, the main research question behind this paper could be formulated as follows: “What is the machine learning

*model that should be used for the online time-series classification of special behavior and how do the different models perform in practice ?”.*

### 3. State of the Art

In this section, we briefly recall the main time-series classification algorithms mentioned in the literature. We compare these techniques based on applicability and effectiveness. In addition to conducting a literature study, we also apply these different techniques to our dataset to test their performance.

For time-series data, there exist several algorithms that consider the time factor, which is essential in our study. A common problematic solution that could happen when dealing with time-series data is to treat each value in the sequence as a separate feature. This is the core difference between time-series data and tabular data. In time-series data, the order of the data is essential and critical. In contrast, in tabular data, the order is ignored and scrambling the order of the features will not affect the prediction process. Therefore, each algorithm dedicated for time-series data is based on a technique and perspective that extracts knowledge from the time-series data concerning the order of the data.

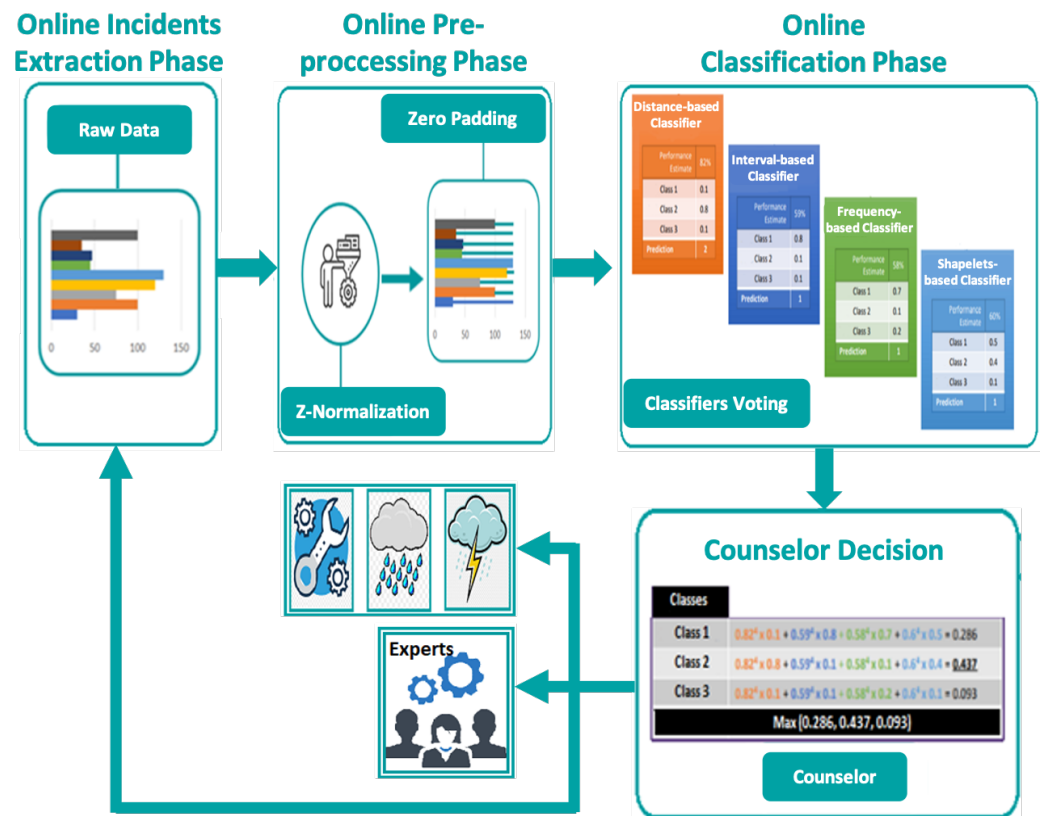
Those algorithms are categorized as follows:

- Distance-based algorithms: This type of algorithms relies on distance metrics to find the optimal class membership. It plays a vital role in pattern recognition problems. The most popular distance measures used are Euclidean [8], Manhattan [9] and Dynamic Time Warping with Barycenter Averaging (DTW) [10] which is the similarity measure used by the K-nearest neighbor algorithm.
- Interval-based algorithms: This algorithm depends—through its classification—on the information retrieved from various series intervals. Time-series forest classifier (TSF) is a classification technique that is built for this type of algorithm [11]. TSF adopts the random forest classifier technique and applies it to time-series data.
- Frequency-based algorithms: classifiers that follow this type of algorithm rely on the frequency of the extracted features from the time-series data. Random interval spectral ensemble known as RISE, is a straightforward classifier that is similar to a time-series forest [12]. Therefore, this algorithm constructs decision trees, and the classification takes place upon the majority of votes.
- Shapelet-based algorithms: the main objective of the shapelet-based algorithm is to identify, for a particular class, the bag of shapelets with discriminatory power. Each shapelet is an interval extracted from a time series and it should follow the same order. During classification, the Shapelet-based algorithm transforms the incoming datasets into “K” shapelets that are yet to be compared by the “K” shapelets extracted for each class in the training phase [13].

In their paper [14], the authors introduced time-series data and time-series classification methods, focusing in their research on the importance of distance-based classifiers. Xing et al. [15], in their paper, divided the time-series classification method into three main categories. Feature-based methods, model-based methods, and distance-based methods. Diving deeply into the literature, we noticed that most of the research works focused on or introduced a specific classification algorithm. As we will see in the experimentation section, when these classification algorithms are applied to our data, they are not able to perform the task in all situations. That is why we propose a novel approach that gives the best results through our testing.

### 4. Online Detection and Prediction Phase

In Figure 3, we depict the three main components of our online detection and prediction phase. It is composed of: (1) the online incidents extraction, (2) the online preprocessing phase, and finally (3) the online classification phase. First, it is important to mention that the intended result is to reduce the errors and not eliminate all errors. As our data are significant and challenging, removing all errors is impossible.



**Figure 3.** The proposed online detection and prediction phase.

The data that are sent by the probes from different locations are continuously monitored. A high reading that is above the peak threshold will trigger the system to check whether the reading of this probe will remain above the peak for 30 min. If the incoming readings remain high (above the peak threshold), this series will be extracted as an incident starting from the value above the maximum background mean until 30 min have passed.

Extracted incidents cannot directly enter into the classification phase. Preprocessing treatment for the raw incidents should be done. Incidents coming from different probes have different characteristics in terms of length, scale and level. During the preprocessing phase, the data of the incidents are normalized using the z-normalization technique and padded using the zero-padding technique. This preprocessing phase will not affect the shape of the incident; it will only standardize the incidents to become similar to the training dataset that the classifiers have trained over. This will help classifiers identify or predict the class of these unknown incidents.

The classification phase is divided into two phases: the voting phase and the counselor's decision. In the voting phase, four classifiers from the state of the art are implemented separately. All of the previous classifiers were implemented and tested because each one was successful in identifying a particular class. Each classifier will accept as input the incoming incident. The four classifiers will run in parallel:

1. The distance-based classifier is implemented using K-nearest neighbor with dynamic time warping (DTW) + barycenter averaging (DBA) as the similarity measure. This classifier can successfully differentiate between calibration and stormy rain classes with an accuracy reaching 89.28%. However, it faces some difficulty separating between rain and stormy rain classes.
2. The frequency-based classifier was built using the random interval spectral ensemble (RISE) algorithm. This algorithm has proved its ability to differentiate between rain class and stormy rain class; its accuracy reached 85.71%. As for the calibration class,

the algorithm had slight errors in classifying calibration incidents as rain and vice versa.

3. The TSF classifier is implemented based on the interval-based algorithm that is similar to the frequency-based algorithm except in the way it slices the series. Each series is split into intervals of varied length within the same decision tree, while RISE performs a random interval length splitting that varies from one decision tree to another but within the same fixed interval. The TSF classifier supports decision making, especially between calibration and rain classes; its accuracy reached 89.28%.
4. The shapelets-based classifier was implemented, although its accuracy was low as it only reached 50%. However, the significance of this algorithm is in differentiating between the calibration class and rain class. This algorithm failed in separating the other classes because of the high similarity in the shape that some incidents of different classes have. Furthermore, as this algorithm creates a bag of Shapelets (sub-shapes of the series) for each class to be used as discriminatory power, confusion may arise.

In our approach, each classifier will perform its prediction and the output will not just be the predicted class but also the probability of each class upon which it concluded to select the class of higher probability. Then, the second phase of the online classification phase is introduced to take all the classifiers' predictions. Its role is to analyze what the majority has classified this new incoming incident as. The counselor has to choose one of three possible choices:

1. The majority of votes of the classifiers and the aggregated probability is high (above 90%). Thus, the decision is directed to assign the incident to this particular class with the highest probability.
2. If the aggregated probability is between 70% and 89%, the collected data are not enough for the counselor to decide. In this case, one will wait for more time to collect additional readings that can help in the decision making.
3. If the probability was low (less than 70%), this means that the incident occurring should be considered as a new incident and a new class of incidents should be created. Here is the role of the experts in the domain to examine this new incident and attempt to identify the nature and the reasons behind this incident. Furthermore, this new incident could be a new shape for an existing class that the model has not been trained on it yet. Thus, in all cases, this new incident is an added value for the model in the future when re-training the classifiers on identifying such cases.

In summary, combining all classifiers' abilities helped overcome the problems and challenges found in our data. The counselor has three possible choices depending on the highest aggregated probability. Suppose that the probability remains high (above 80%) after three classification attempts. In that case, the incident will be assigned to the class of the highest probability. If the probability varies between 60% and 80%, then the incident will undergo further classification after collecting more incoming data readings. Moreover, the incident is left for the experts to check and verify whether it did not succeed in gaining a probability higher than 60% so that it could be a new incident of a new class to be created or a new shape for the existing class.

## 5. Experimentation

In order to compare the different approaches of the state of the art, as well as to see the benefit of our proposed model, we decided to evaluate systematically different experiments and evaluations on the labeled time-series data. After investigating all of the mentioned algorithms in the state of the art, we attempted to implement each classifier based on its best practice for selecting the optimal parameters and then apply it over our labeled data. First, the data we have are split into two sets, a training dataset (90% of the dataset) and a testing dataset (forms the rest 10%). When splitting the data between training and testing, we guarantee that the training dataset is balanced and presented well in all three classes so that the classifier will be trained well.



All the implementations are performed using Python libraries. For the best environment performance and for easy implementation, we installed Anaconda, in which we used the Jupyter notebook for writing and testing the code. Anaconda provides us with an isolated environment containing all the needed libraries to perform our tests. Going deeply into the libraries dedicated to time-series data, several methods are defined to handle this type of data. The traditional machine learning algorithms implemented for tabular classifiers cannot be applied in our case of time-series data because these neglect the time factor essential in our data. Thus, in addition to Sklearn, pandas, numpy, and other libraries, we installed and used the Sktime library which contains the time-series classifiers.

In Table 1, we found the evaluation results for the four times series classifiers of the state of the art. Class 0 corresponds to the calibration cluster, Class 1 to the rain class and class 2 to the stormy rain class.

The first classifier is the KNN with DTW, which is a distance-based classifier. By default, this classifier uses the Euclidean distance measure [8] to determine the membership of a class. For our case, the time-series data require a different metric algorithm because incidents are of varying length and are not perfectly aligned in time. Although the accuracy was not bad (89.28%), after training and testing it, some errors still occurred. By investigating what the model failed, we deduced that it could detect the calibration class and the rain class but failed to identify the stormy rain class. The model got confused between the rain class and stormy rain class incidents and classified the stormy rain as rain incidents.

The second classifier is the time-series forest classifier which relies on the interval-based algorithm. This classifier depends on the information retrieved from the various intervals of a series. At first, the classifier splits the series into random intervals; each has a random starting point and length. Then, the algorithm extracts summary features (slope, mean, and standard deviation) from those intervals. The extracted features form the feature vector representing the interval. Since this algorithm is based on the random forest algorithm, it will construct and train a decision tree from the extracted features. Several trees are constructed to support decision making and select the majority of the trees in the forest. After training and testing the TSF classification model, the performance was good (89.3%), but not sufficient. The model was able to identify the calibration class but it faced some errors when identifying the stormy rain class.

**Table 1.** Applying different classification algorithms to the dataset.

Classification Algorithm	Accuracy	Distinguished Classes	
		Pass	Fail
KNN+DTW	89.28%	Class 0 & Class 1 Class 0 & Class 2	Class 1 and Class 2
TSF	89.3%	Class 0 and Class 2	Class 0 & Class 1 Class 1 & Class 2
RISE	85.71%	Class 1 and Class 2	Class 0 and Class 1 Class 0 and Class 2
Shapelet	50%	Class 0 and Class 2	Class 0 & Class 1 Class 1 and Class 2

The third classifier is the random interval spectral ensemble (RISE) classifier. This classifier is based on the frequency features extracted from the series after splitting it into intervals. It sounds similar to the previous classifier, the TSF, especially because it also uses the random forest algorithm. It differs from TSF in two ways. First is how it splits the series into intervals, where the intervals for each decision tree are of the same length. The second difference is in the type of features that the algorithm extracts from the intervals, where RISE extracts spectral features (series-to-series features) and not summary statistics. The algorithm was significant in classifying the rain class from the data. In the rest of the classes, however, it faced some errors. The accuracy of this model reached (85.71%).

Finally, the last classifier is the shapelet-based classifier. This classifier is very popular and used when dealing with time-series data. A shapelet is a sub-shape of a series. A bag

of shapelets is used to represent a particular class. When extracting those shapelets, the algorithm searches for shapes with discriminatory power to identify a class. Shapelets form the identity of each class. When a new unknown incident arrives, the algorithm will extract its shapelets and compare them to the classes' shapelets to confirm which class the incident belongs to. The shapelet-based algorithm was implemented and tested over our data, but the results were unsatisfying. After several tests and attempts to enhance the model's overall accuracy, it nearly reached 50%. However, after we investigated the results, we uncovered the reason for such an outcome. The data that we have are very challenging because they are very similar to each other, which makes their shapes very similar; this is why the model was confused. Even though the classification model was able to identify the rain class, it failed in the other two classes (calibration and stormy rain).

To start evaluating our online classification module, we first tested incoming incidents. These incidents are preprocessed online after being extracted and then prepared to be classified with the classification algorithms. Then, the classification algorithms will work individually in parallel on the incoming incident, trying to classify it as soon as possible. The classification algorithms presented in Table 2 will return their predictions for the incoming incidents as soon as possible. This prediction will be in the form of a probability suggested by each algorithm to each incident while trying to map it to the respective class.

**Table 2.** Testing our counselor approach.

	KNN+DTW	TSF	RISE	Shapelet	Counselor Performance	Counselor Decision
<b>Incident 1</b>	C0 (100%)	C0 (95%)	C0 (63%)	C0 (34%)	C0 (73%)	Wait for more data
	C1 (0%)	C1 (0%)	C1 (6%)	C1 (21%)	C1 (6.75%)	
	C2 (0%)	C2 (5%)	C2 (31%)	C2 (45%)	C2 (20.25%)	
<b>Incident 2</b>	C0 (98%)	C0 (96%)	C0 (68%)	C0 (99%)	C0 (90.25%)	C0
	C1 (2%)	C1 (1%)	C1 (23%)	C1 (0%)	C1 (6.5%)	
	C2 (0%)	C2 (3%)	C2 (9%)	C2 (1%)	C2 (3.25%)	
<b>Incident 3</b>	C0 (0%)	C0 (0%)	C0 (0%)	C0 (0%)	C0 (0%)	C2
	C1 (31%)	C1 (0%)	C1 (3%)	C1 (0%)	C1 (8.5%)	
	C2 (69%)	C2 (100%)	C2 (97%)	C2 (100%)	C2 (91.5%)	

Finally, the counselor will start performing the task assigned to it. Thus, the role of the counselor will be to decide which algorithm acts the best and gives the perfect prediction for the incoming incident, as shown in Table 2, where incidents 2 and 3 were assigned to classes 0 and 2, respectively. However, incident 1's probability was not enough for the counselor to make a decision, which it is it suggested waiting for more data.

Our proposed model overcomes the issue that we were concerned about. When tested on its own, the problem with each classification algorithm was its ability to identify a single class and failing in differentiating between the rest of the classes. By combining the outputs of those four algorithms, the counselor was able to either commit the identity of the unknown incoming incident or consider it a new incident related to a new class to be examined by experts. Therefore, the proposed classification model output was satisfying and it supported decision making for predicting the class of incoming incidents.

## 6. Conclusions

In this work, we presented our machine learning-based framework for autonomously identifying the causes behind the online incoming incidents caused by high gamma dose rate readings. After extracting, preprocessing, and clustering the historical incidents, our approach is to apply a machine learning model that will match online incoming incidents to their similar clustered ones to identify the causes behind them as soon as possible using supervised classification.

In the classification phase, we, specifically, tackled the problem of classifying time series using several classification algorithms at the same time which was properly addressed



nowhere in the literature. We researched and experimented with the different classic and state-of-the-art approaches to evaluate their compatibility. When those approaches failed to classify our data when properly testing each approach alone, we proposed our counselor classification model for using all the classification algorithms simultaneously and voting for the one with the best outcome.

Displaying the obtained matching percentages through the experimentation of various algorithms, we were able to highlight how our model comparatively gave the best results. Furthermore, the experts expressed positive and hopeful thoughts upon inspecting the results, which motivated us to publish this contribution in an article. As future work, the next step would be to improve the quality of the overall framework by exploring the evaluation with more datasets to automate the evaluation as well.

**Author Contributions:** M.A.S. analyzed and interpreted the need for an intelligent radiation early warning system, introduced the second phase of the RIMI framework, and was a major contributor in writing the manuscript. He also queried the AI methodologies and techniques to introduce an approach that can address the shortcomings behind the RIMI second phase. B.F., Y.T., A.J. and R.L. verified the tested techniques and functions for analyzing the data and were major contributors in writing the manuscript. All authors read and agreed to the published version of the manuscript.

**Funding:** The first and corresponding author “Mohammed Al Saleh” has a scholarship from the National Council for Scientific Research in Lebanon (CNRS) to continue his Ph.D. degree, including this research.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

**Acknowledgments:** We would like to thank the National Council for Scientific Research (CNRS) in Lebanon for supporting this work. We would like to express our gratitude to the Federal Office for Radiation Protection (BfS) in Germany for allowing us to use the data collected by their REWS since more than 15 years ago. We would also like to thank Roy Issa for his support in implementing the code behind this research.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

IoT	Internet of Things
REWS	Radiation Early Warning System
DTW	Dynamic Time Warping
DBA	DTW Barycenter Averaging
KNN	k-Nearest Neighbors
TSF	Time-Series Forest
RISE	Random Interval Spectral Ensemble

## References

1. Mann, W.B. The international chernobyl project technical report: Assessment of radiological consequences and evaluation of protective measures. *Appl. Radiat. Isot.* **1993**, *44*, 985–988. [[CrossRef](#)]
2. Thieu, D.; Toan, T.N.; My, N.; Sy, N.; Tien, V.; Mai, N.; Cuong, L. Study, Design and Construction of an Early Warning Environmental Radiation Monitoring Station. *Commun. Phys.* **2012**, *22*, 375–382. [[CrossRef](#)]
3. Stöhlker, U.; Bleher, M.; Doll, H.; Dombrowski, H.; Harms, W.; Hellmann, I.; Luff, R.; Prommer, B.; Seifert, S.; Weiler, F. The German Dose Rate Monitoring Network Furthermore, Implemented Data Harmonization Techniques. *Radiat. Prot. Dosim.* **2019**, *183*, 405–417. [[CrossRef](#)] [[PubMed](#)]
4. Al-Saleh, M.; Finance, B.; Haque, R.; Taher, Y.; Jaber, A. *Towards an Autonomous Radiation Early Warning System*; BDCSIntell: Versailles, France, 2019.

5. Z-Normalization of Time Series. Available online: [https://jmotif.github.io/sax-vsm\\_site/morea/algorithm/znorm.html](https://jmotif.github.io/sax-vsm_site/morea/algorithm/znorm.html) (accessed on 27 January 2022).
6. Myers, C.S.; Rabiner, L.R. Connected digit recognition using a level-building DTW algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **1981**, *29*, 351–363. [[CrossRef](#)]
7. MacQueen, J. Some methods for classification and analysis of multivariate observations. *Comput. Chem.* **1967**, *4*, 281–297.
8. Gower, J.C. Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra Appl.* **1985**, *67*, 81–97. [[CrossRef](#)]
9. Suwanda, R.; Syahputra, Z.; Zamzami, E.M. Analysis of Euclidean Distance and Manhattan Distance in the K-Means Algorithm for Variations Number of Centroid K. *J. Phys. Conf. Ser.* **2020**, *1566*. [[CrossRef](#)]
10. Anh, D.T.; Thanh, L. An efficient implementation of k-means clustering for time series data with DTW distance. *Int. J. Bus. Intell. Data Min.* **2015**, *10*, 213–232. [[CrossRef](#)]
11. Deng, H.; Runger, G.; Tuv, E.; Vladimir, M. A Time Series Forest for Classification and Feature Extraction. *Inf. Sci.* **2013**, *239*, 142–153. [[CrossRef](#)]
12. Flynn, M.; Large, J.; Bagnall, T. *The Contract Random Interval Spectral Ensemble (c-RISE): The Effect of Contracting a Classifier on Accuracy*; Hybrid Artificial Intelligent Systems; Springer: Cham, Switzerland, 2019.
13. Zhang, J.; Shen, W.; Gao, L.; Li, X.; Wen, L. Time Series Classification by Shapelet Dictionary Learning with SVM-Based Ensemble Classifier. *Comput. Intell. Neurosci.* **2021**, *2021*, 5586273. [[CrossRef](#)]
14. Abanda, A.; Mori, U.; Lozano, J. A review on distance based time series classification. *Data Min. Knowl. Discov.* **2019**, *33*, 378–412. [[CrossRef](#)]
15. Xing, Z.Z.; Pei, J.; Keogh, K. A Brief Survey on Sequence Classification. *SIGKDD Explor.* **2010**, *12*, 40–48. [[CrossRef](#)]