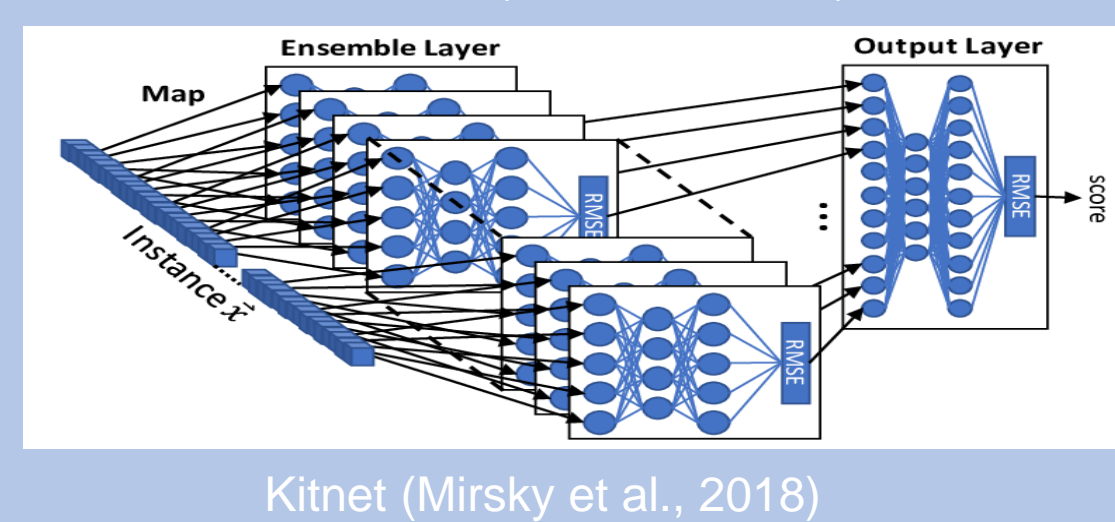
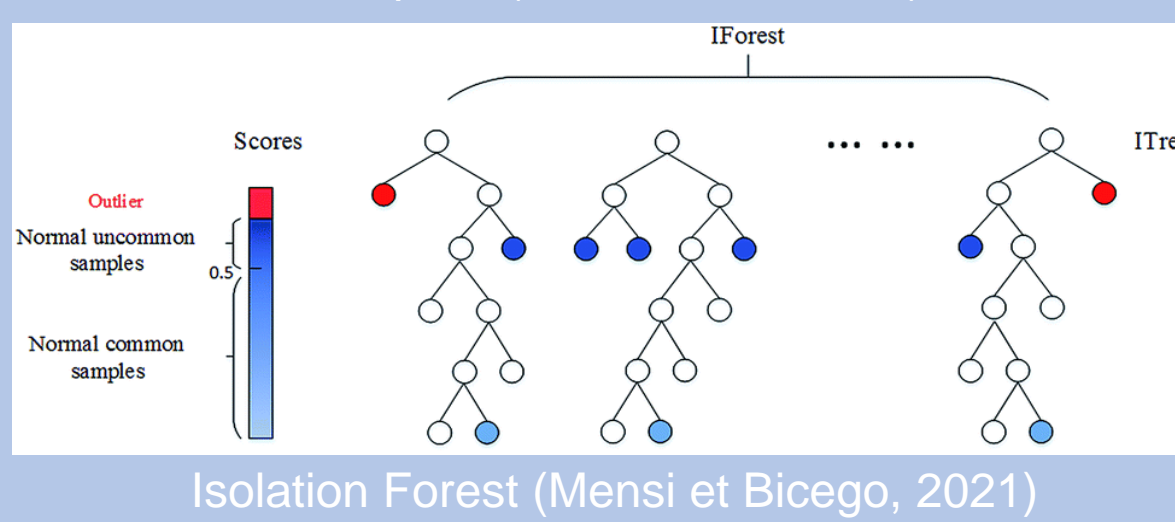
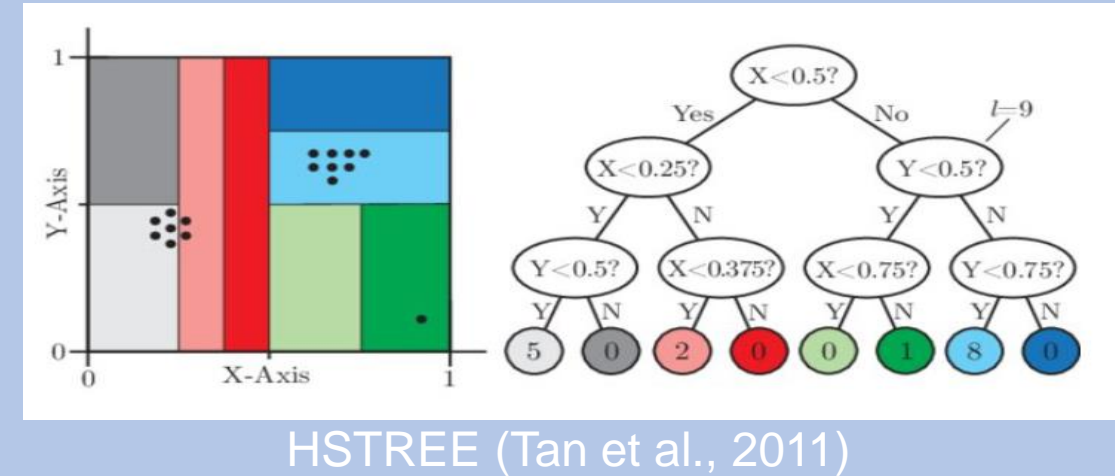
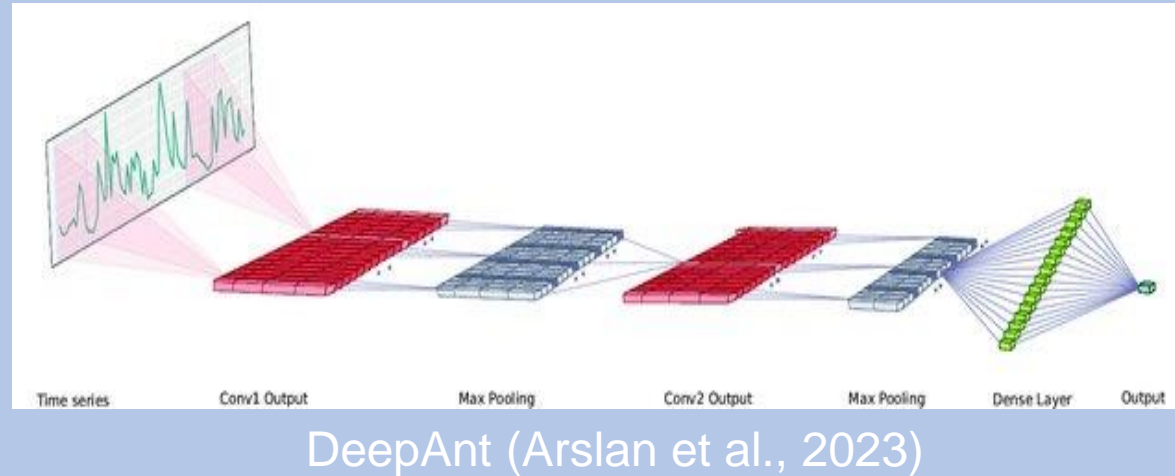


## ABSTRACT

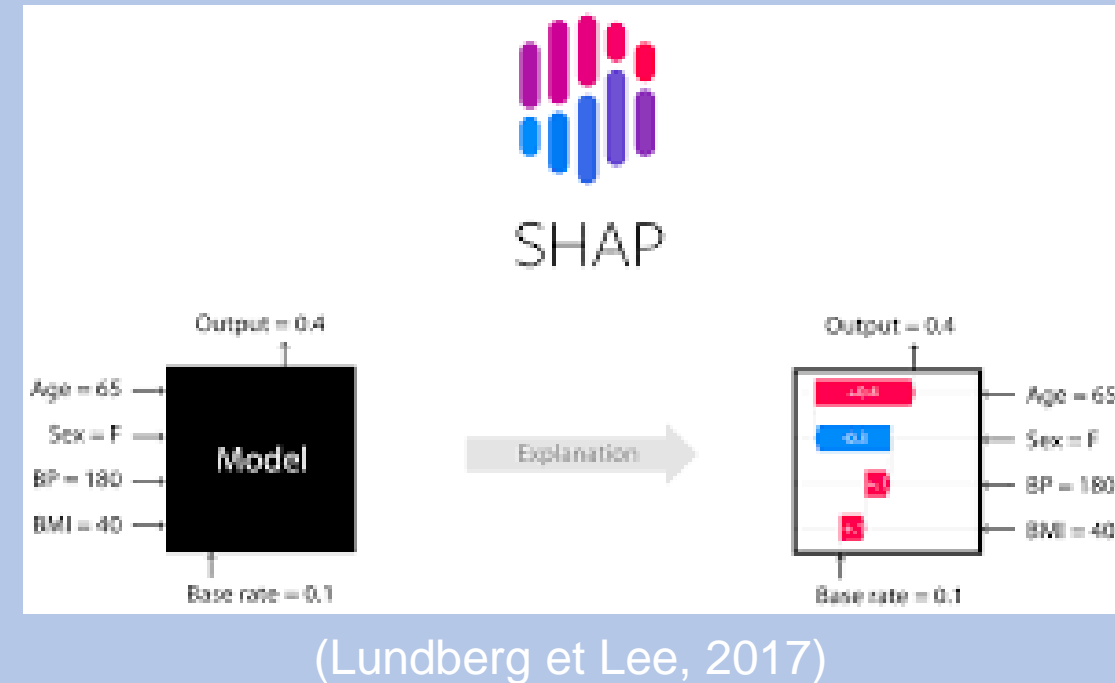
Anomaly detection refers to the identification of rare events that differ significantly from the normal trend observed in the data distribution. When the number of variables to analyze is large, it can be difficult to understand the detected anomaly without explanation. In this work, we present the prototype of an explainable real-time anomaly detection system, based on measurements from a multivariate data stream which can be assimilated to an infinite multivariate time series. The built system is composed of a set of anomaly detection methods combining deep neural networks and decision trees as well as an agnostic explainability method. In an unsupervised learning context, we also show how explainability provides insights to validate the system.

## MODELS

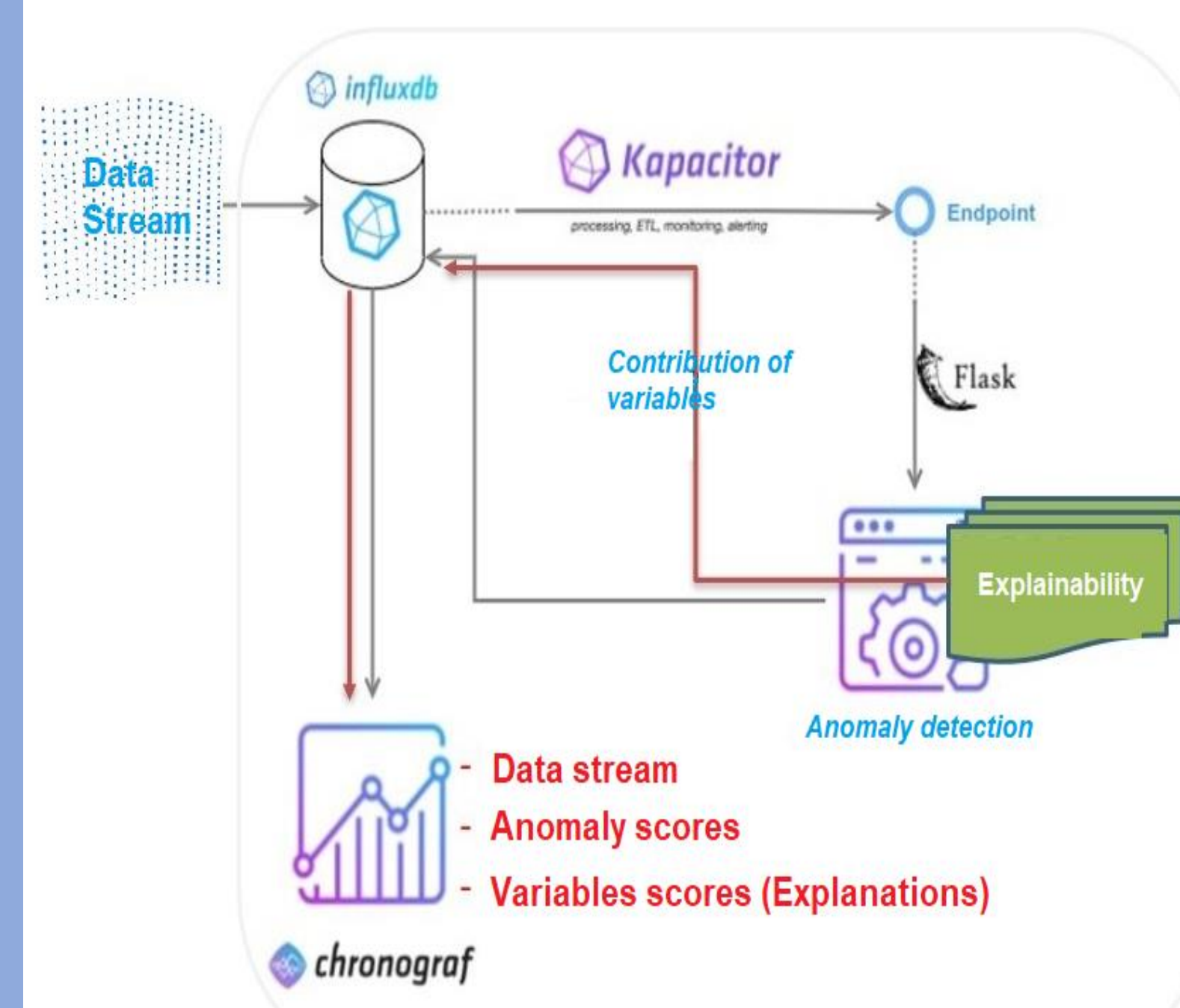
Some anomaly detection models based on different approaches was preliminary integrated to showcase the overall system idea.



- Explanability provides insights to understand the anomaly.
- In the context of multivariate data, a score attribution explainability method can help understand which variables contributed the most to the anomaly score.



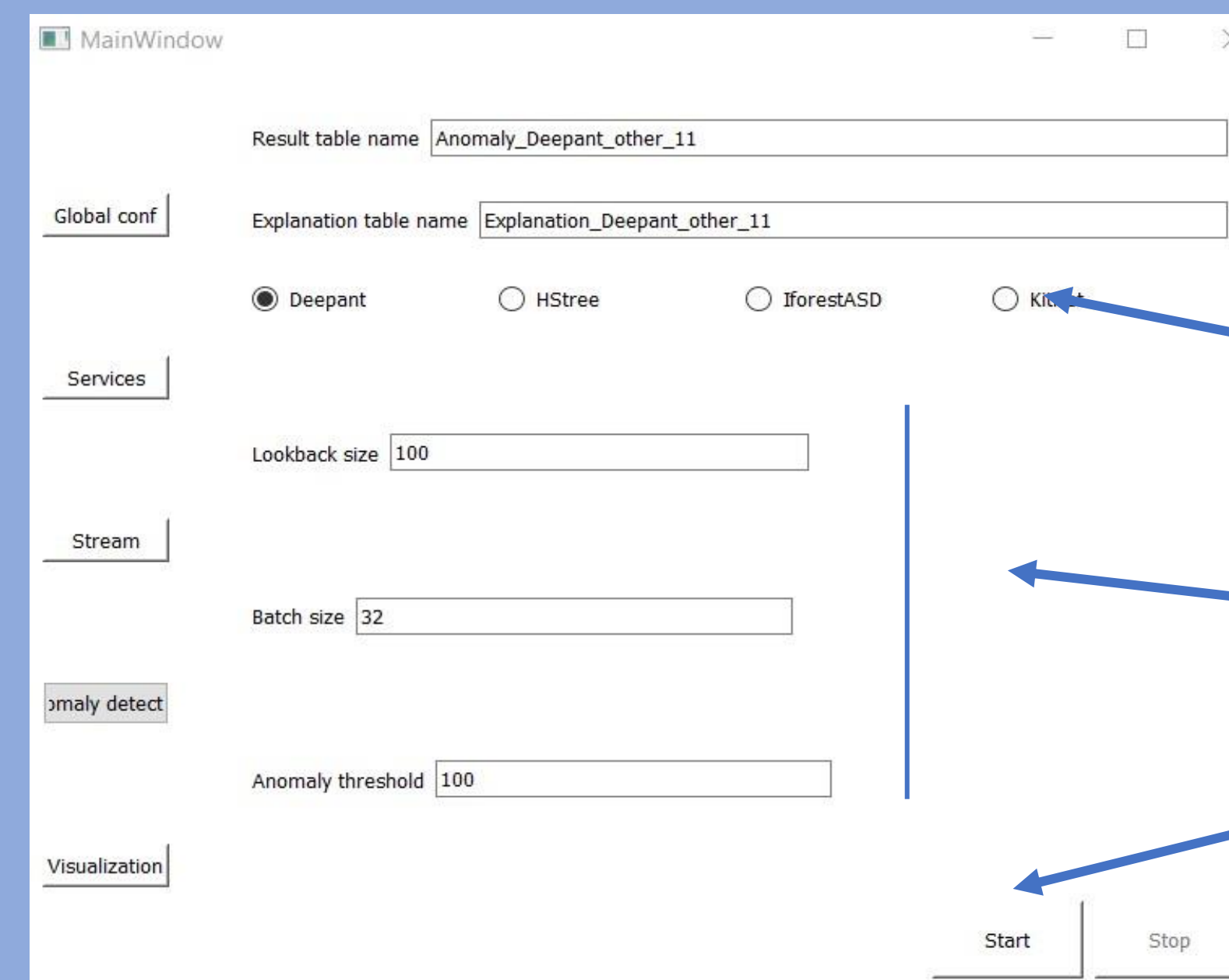
## SYSTEM ARCHITECTURE



The system consists of four main components:

- The time series database stores the data stream, the anomaly scores and the contributions of variables.
- The anomaly detection module analyse and compute the anomaly score for each incoming data point.
- The explanation method computes a score for each variable that denotes its contribution to the anomaly score.
- The visualisation tool shows on the same dashboard, the stream, the anomaly scores and the explanations

## USER INTERFACE TO CONFIGURE AND LAUNCH THE SYSTEM



The GUI enables the configuration of the services and the anomaly detection algorithms.

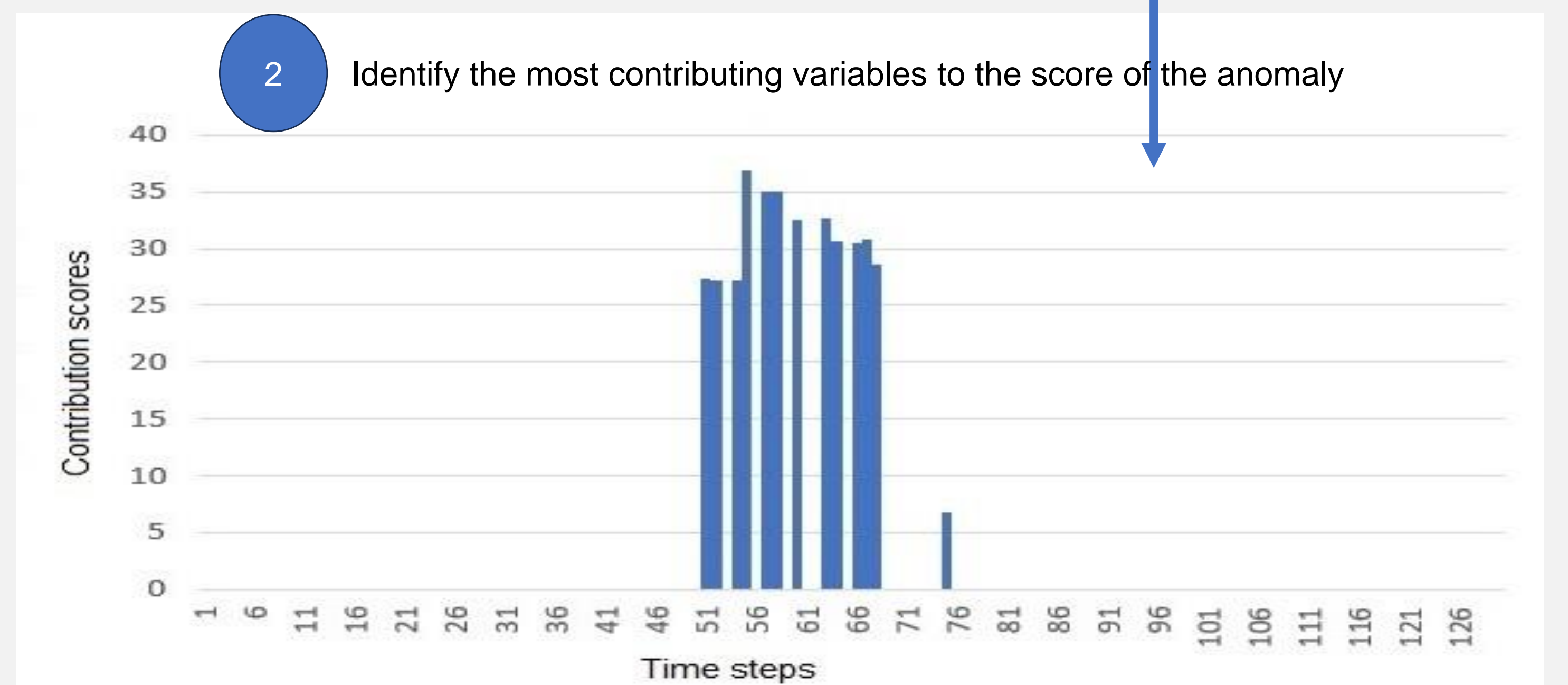
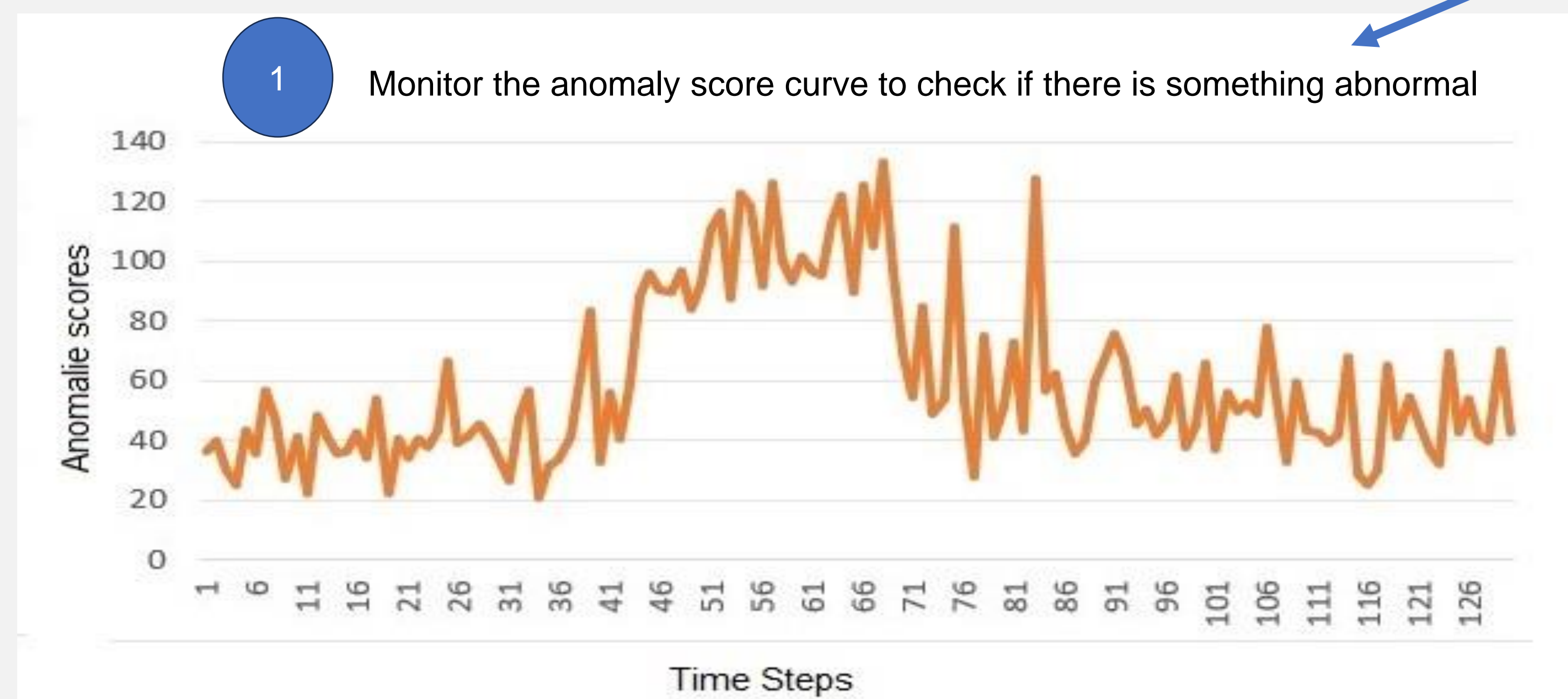
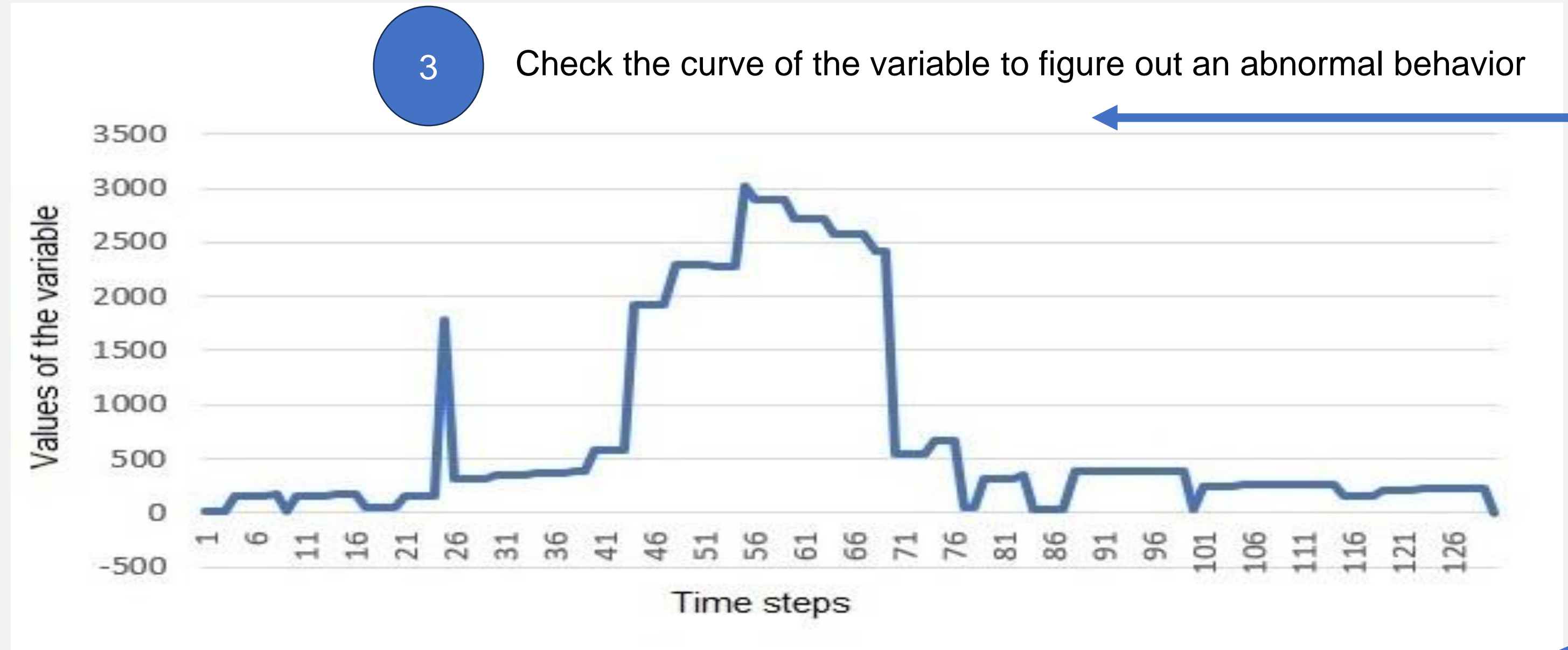
Choose the anomaly detection algorithm

Define the parameters of the algorithm

Start the analysis

## VALIDATION

The system implemented has been used to detect and explain anomalies from a multivariate data stream simulated from real data contained in a CSV file. The validation mainly consisted in relying on the explanations provided by the explainability method to get the variables with the highest contribution. Then, observe the curve of those variables at the same time to figure out the anomaly.



## FUTURE DIRECTIONS

- Design a continuous learning model that will be able to handle the concept drift effectively
- Design an accurate explainability method that will be able to provide real time explanations
- Improve the system design and implementation
- Improve the user interface

## ACKNOWLEDGMENTS

DASMA Project benefited from the funding and support from the following entities :

- Banque Publique d'Investissement Bpifrance
- Isite CAP 20 - 25 (<https://cap2025.fr/le-projet-i-site>)
- Ministère de l'Enseignement Supérieur et de la Recherche (MESR).

## REFERENCES

- Florentin Jiechieu et al. SEDAF : Prototype d'un Système Explicable de Détection d'Anomalies dans les Flux de Données, EGC 2024: xxx-xxx, Dijon, Janvier. RNTI.
- Ngo Bibinbe et al. (2022b). A survey on unsupervised learning algorithms for detecting abnormal points in streaming data. In Intl. Joint Conf. on Neural Networks, IJCNN, Padua, Italy, July 18-23, pp. 1-8.
- Lundberg, S. M. et S. Lee (2017). A unified approach to interpreting model predictions. In Proc. of the 31st Intl. Conf. on Neural Information Processing Systems, NIPS, Red Hook, NY, USA, pp. 4768-4777. Curran Associates Inc.