# 1 - Introduction

The targets imposed by the European community impose the reduction of consumption by 2030 and in this perspective the reduction of energy consumption in buildings becomes of fundamental importance. It is estimated that most of the consumption comes from the operation 90% of the total energy consumed during the life cycle of a building.

The reduction of energy usage in building could support the urgency to reduce world emissions, though Promote energy awareness and prevent wastes.

a large amount of data is produced by the operation of buildings thanks to the increasingly widespread introduction of IOT devices, reaching zettabyte \cite{ ﻿Erhan2021 }. which have made buildings become not only and energy intesndive but information intensive \cite{﻿Fan2021}. Smart cognitive building \cite{﻿Rinaldi2020}

Building data are heterogeneous from occupancy patterns to ambient environmental conditions to appliance parameters and energy consumption. All these data each of these data reflects a complex interaction that occurs between occupants, energy systems, the building envelope, and external conditions. Managing those data is not trivial \cite{ ﻿Molina-Solana2017 }

The availability of massive building operational data, if properly managed and analyzed, provide the opportunity to gain insight on the building operational behaviour discovering opportunities for savings. Promoting sustainability behaviour.

Machine learning Data driven approaches can be helpful to variety of tasks: pattern recognition, analysis for performing essential tasks in building energy management such as energy consumption forecasting, anomaly detection and diagnosis, advanced benchmarking, load profiling, and schedule optimization of building energy systems. \cite{ ﻿Molina-Solana2017 }

Anomaly detection consistes in detecting patterns that eviates from expected behaviour, can be employed in detecting abnormal behaviour of end users, detection of faulty appliance or subsystem

Recent studies have been demonstrated that the use of artificial intelligence and effective data analytics techniquest

## 1.1 - General definition of anomaly

Ana anomaly is a point that deviates from normal expected behaviour:

A point anomaly means that one individual event instance can be considered anomalous when compared to the remaining data.

Context anomalies start from the assumption of dividing the behaviour from the context: the same behaviour might not be considered an anomaly if it happens in a different context.

collective anomaly. In this case, the event instance does not represent an anomaly per se, but only if considered within the collection of all the other events instances.

## 1.2 – Related work for anomaly detection

\cite{Himeur2020} performs a classification of anomaly detection techniques on algorithmic centric view in building energy systems

Detection level

* Aggregated level
* Appliance level
* Spatiotemporal level

Algorithm

* Supervised

Unsupervised

\cite{﻿Voltage2016} performs anomaly detection on smart grid though the use of clustering

* Feature extraction
* This is the case of distance based techniques of MP

## 1.3 - Anomaly detection using Matrix Profile

is at aggregated level unsupervised feature extraction

One of the most promising technique is Matrix profile . Firstly introduced by \cite{} it is …

The main properties of this methods are…

* It is exact: the Matrix Profile based methods provide no false positives or false dismissals. It can handle missing data: Even in the presence of missing data, we can provide answers which are guaranteed to have no false negatives.
* It is simple and parameter-free: In contrast, the more general algorithms in this space that typically require building and tuning spatial access methods and/or hash functions.
* It is space efficient: Matrix Profile construction algorithms requires an inconsequential space overhead, just linear in the time series length with a small constant factor, allowing massive datasets to be processed in main memory (for most data mining, disk is death).
* It is incrementally maintainable: Having computed the Matrix Profile for a dataset, we can incrementally update it very efficiently. In many domains this means we can effectively maintain exact joins/motifs/discords on streaming data forever. MP is extremely scalable, for extremely large datasets we can compute the Matrix Profile in an anytime fashion, allowing ultra-fast approximate solutions and real-time data interaction.
* Simplicity and Intuitiveness: Seeing the world through the MP lens often invites/suggests simple and elegant solutions.
* It can be constructed in deterministic time: given only the length of the time series, we can precisely predict in advance how long it will take to compute the Matrix Profile. (this allows resource planning)
* It can leverage hardware: Matrix Profile construction is embarrassingly parallelizable, both on multicore processors, GPUs, distributed systems etc.

Matrix profile has been used for anomaly detection I different fields

informatics

\cite{DinalHerath2019} introduces ﻿a real time anomaly detection framework based on matrix profile called RAMP (Real-Time Aggregated Matrix Profile), that is able to identify anomalies in scientific workflows. (building block).

An industrial application of anomaly detection is presented in\cite{﻿Anton2020} which combines the classical approach of Matrix profile with the hamming distance to automatically detect intrusions in the network of a water processing facility.

\cite{ ﻿DePaepe2020a} Applies a noise elimination technique based on Matrix Profile on real Yahoo! internet traffic metrics to detect anomalous behaviours

\cite{ ﻿DePaepe2019 } demonstrate how the elimination of noise can help in anomaly detection of noisy date by testing the algorithm on numenta benchmark

\cite{ ﻿Madrid2019 } applies the pan matrix profile algorithm to find different length anomalies in ﻿automated pedestrian counting system developed in Taipei

*medicine*

\cite{﻿Alshaer2020} proposes an unsupervised real time anomaly detection method based on continuous learning of time series shaplets extracted though Matrix Profile algorithm. Those shaplets are extracted and stored in an anomaly library and then used for anomaly detection in an electro-cardiogram (ECG) time series (﻿MIT-BIH database [32]), using a in a sliding window.

*energy*

﻿\cite{ ﻿Nichiforov2020 } Identifies anomalous patterns though a basic application of Matrix Profile on public building energy traces and then classifies the pattern

\cite{ ﻿Zhu2020} demonstrates how Matrix Profile can be useful in detecting anomalies in different fields in particular in meter swapping and earthquake monitoring.

\cite{﻿DePaepe2020b} applied an implementation of the classic Matrix Profile, called Contextual Matrix Profile, in detection of anomalous energy consumption on a ventilation units of three households.

## 1.4 - Problem of classic approach and improvements

Introduction of domain knowledge

﻿\cite{Dau2017} introduces the concept of annotation vector used to introduce domain knowledge in the process of motif and discord discovery, which allows to find results that follows users defined constraint and produce better results, closer to expectations of the analyst. This method has been proved to be effective to solve different issues: simplicity bias, actionability bias.

However this method is a posteriori method that does not modify the way MP is calculated

… contextual matrix prefile

Twin freak

﻿For a given sub-sequence, Matrix Profile computes the Euclidean distance with respect to all other sub- sequences and identifies the minimum distance. Therefore, a repeated anomaly instance would cause false negatives due to the previous anomaly instance being part of the all sub- sequence set.

Specifically, frequent/rare sub sequences are defined as the ones with the smallest/largest 1-nearest neighbour distance, which are also known as motif/discord. However, discord fails

the ones with the smallest/largest 1-nearest neighbour distance, which are also known as motif/discord. However, discord fails to identify rare sub sequences when it occurs more than once in the time series, which is widely known as the twin freak problem.

\cite{DinalHerath2019} through a semi-supervised model permits to limits the number of sub sequences compared, considering for comparison only references with no anomalies.

\cite{﻿He2020} proposes a method called “Neighbour Profile” based on sampling and density estimation to perform anomaly detection and overcame the issue of twin freak.

# 2.- Contribution of the paper

# 3.-Description of the Data Analysis Methods

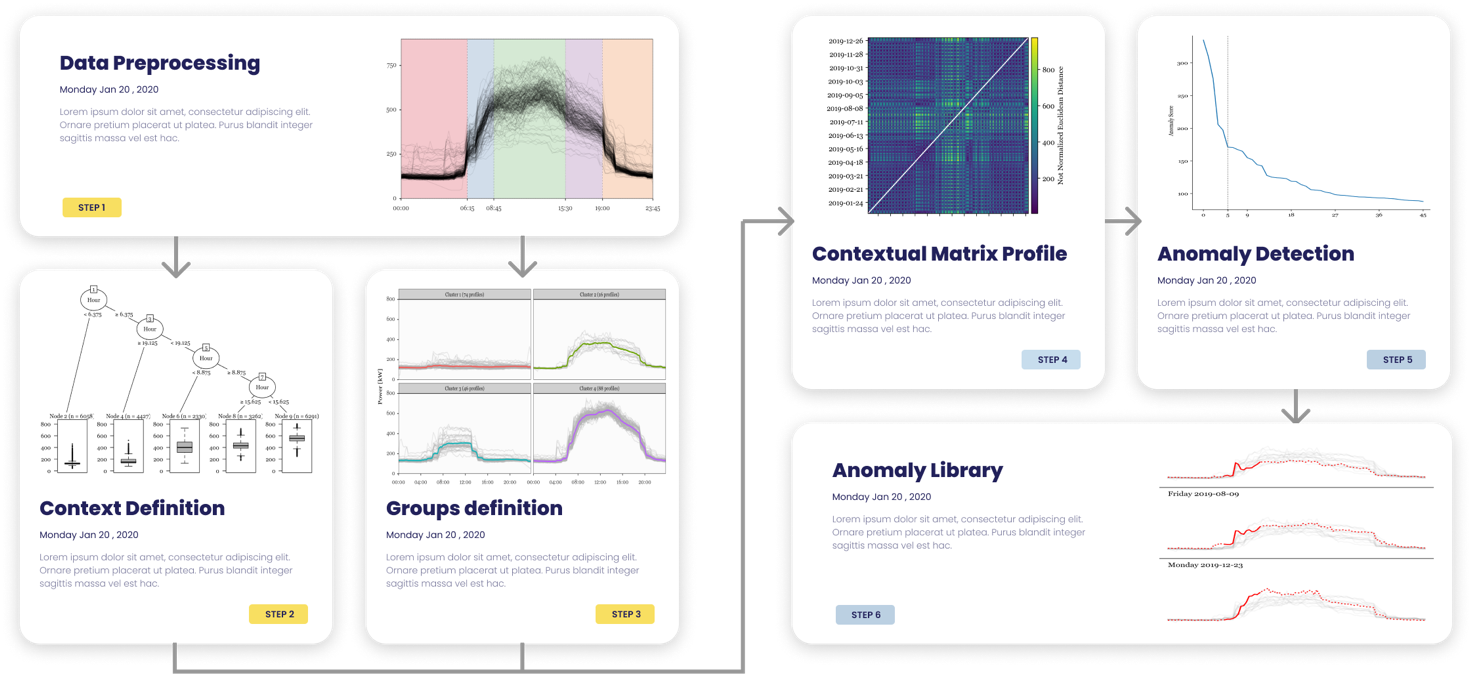
Contextual Matrix Profile

Anomaly score

CART

classification

# 4. - Methodological Framework



Framework generale e pseudocodice?

# 5. - Case study

The case study analysed refers to the energy consumption of a MV/LV transformer cabin identified as “substation C”, that serves a part of the main campus of Politecnico di Torino (PoliTo), an Italian university located in Turin. The measurement infrastructure provides the total electrical load with 15 min timestamps. In order to use a dataset that is large enough to capture the behaviour of the electrical load, with regular occupation patterns

# 6. - Results

# 7. - Discussion

# 8. - Conclusions and Future Work