AD! methods can be categorized in many ways [6]. In this paper is divided in two parts (cf Figure 1), general methods and useful tricks to improve theses. The former represents general ways to solve the **AD!** problem. It is composed of three categorizes: forecasting (Sec 0.1), clustering (Sec 0.2) and index monitoring (Sec. 0.3). The latter are so-called add-ons (Sec 0.4). The papers in this category present an **AD!** method with some tricks that, in our minds can be applied to any method to improve it.

0.1 Forecasting Methods

When working with some equipment, you know it so well that, if asked, you could almost predict the sound it will make in the near future. But if suddenly the sound doesn't match your expectations you know there are some things wrong. That is the core idea for the forecasting method, predict the near future using some generative method and if the error goes over some threshold, the data point is labeled as an anomaly. Example of those forecasting technics may be **LSTM!** [9] or using **TCN!** [1].

0.2 Clustering method

While storing metric screws in the warehouse, someone picks up an imperial one, he wants to find any box to put it in. Therefore he will conclude there was an error in the shipment, even if he didn't know this type of screw. This is the main idea for the clustering method, define clusters, if a new data point can't fit in one of them, it is labeled as an anomaly. There are multiple ways to define the cluster. A good part uses some defined distance [7] function and uses a variant of **KNN!** (**KNN!**) [12]. Similarly, use the density of the data point in relation to the training data, which will give less, strict categorizes [3]. Another method is to estimate the probability density function (PDF) and put a threshold on low values [2].

0.3 Index monitoring

After assembly, the disks cutters there are put under pressure for some time. When the time is up, the pressure is measured, if the pressure is too low, there is a problem somewhere. This is the intuition behind index monitoring method, define some function with an acceptable range of normal values, if the values are outside the range it's marked as an anomaly. This technic is one of the oldest, simplest and most used in the manufacturing industries, in this context is referred as Control chart [14]. Another example of application is [10] which use a modified wavelet transform to define a Health Index to monitor bearings.

0.4 Add-ones

In order to improve those techniques and solve issues intrinsic to the TBM case, here are some interesting solutions.
Topology, TBM AD is intrinsically a multivariate problem (ie there is more than one sensor on the machine). And it is safe to assume that the anomaly information for some problems is carried by multiple sensors. For example, if there is a defective sensor, the others want to raise any anomaly, which will greatly help to pin the root cause down. This what [8] and [4] did so by introducing a directed graph to model the relationship between sensors. This information was then used, in these papers, to do forecasting using attention techniques.

Contamination, the training data for those models will surely be some already-bore tunnel data. During those bores, by Murphy's law, at some point, some things went wrong. The issue is that most methods assume an anomaly free dataset. This is why [11] [5] [15] proposed some modifications to the training steps to deal with this contamination. To do so, they generally modify the loss function to account for the uncertainty of the data and improve this uncertainty iteratively.

Dimension reduction, a large quantity of sensors, with a quite small data rate, even with a small time window can make the input dimension of the model certainly significant. This can render the training hard and sensitive to noise. to solve this issue, most papers use some kind of dimension reduction. For example, [2] uses an AE and a modified reconstruction error concatenated with the latent vector as an input to a GMM. Or [13] used a **LSTM!** as an AE to do forecasting.

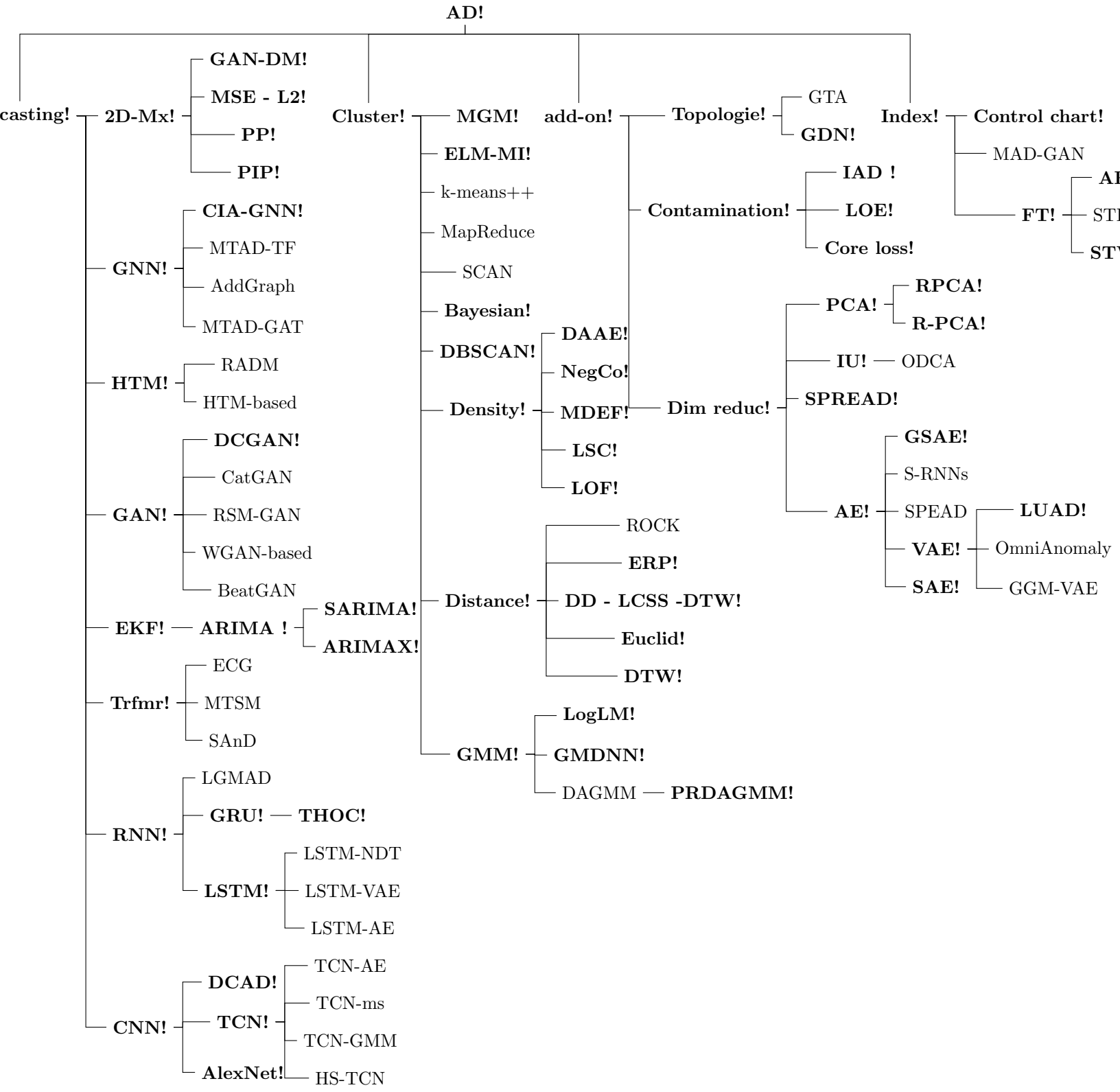

Figure 1: Anomaly detection method categorization