Identifying different operational regimes of an Electric Arc Furnace using Machine Learning based alternative clustering approaches

Venkata Chaitanya Kanakamedala*, Allen Chacko Johny*, Arya K. Bhattacharya*, Kumar Chatterjee†
* Mahindra University, Hyderabad, India
† MCARTech, Inc., Indiana, USA

Abstract— **Machine Learning has been playing a significant and growing role in Industry 4.0. While the most typical application is associated with condition monitoring of equipment and processes by predicting their behaviors into the near future, a more abstract and profound application lies in identifying the operating regimes of complex industrial processes. Industrial processes have different operating modes associated with varying levels of productivity, stability and product quality which can be called as operational regimes, and on certain occasions shift from one regime to another. Automatic identification of operating regimes is a challenge before Machine Learning, and potentially Unsupervised Learning, specifically Clustering, can be used to identify these. In this work, Fuzzy C-Means Clustering has been used to successfully achieve fine distinction between different regimes of an Electric Arc Furnace in the steelmaking process chain. A surprising observation is that while the more popular DBSCAN approach works in distinguishing between gross variations in operation, in finer cases with noisy industrial data, it fails to synthesize realistic distinctions. However, Fuzzy C-means is found to work well and provides realistic insights to plant operations.**

Keywords—EAF, ID-fan, ALERT, Fuzzy C-Means, DBSCAN

I. INTRODUCTION

Industry 4.0 manifests the current state of the evolution of cyber-physical systems towards applications in process and manufacturing industry. It involves incorporating Internet of Things (IoT) applications to improve manufacturing, distribution, and analytics among other aspects. This research work aims to provide a Machine Learning (ML) based solution to a specific problem in an industry to partially enable its shift towards Industry 4.0.

Steel manufacturing is an important segment of the metal industry. A steel manufacturing plant has many crucial components that usually work sequentially like links in a chain to produce the required products. A failure in one of these components could have a ripple effect leading to a total breakdown of the plant. Thus, it becomes necessary to reduce the possibility of such failures. This is where industry 4.0 plays an expanding role.

Javaid et al in [1] provide a survey of various Artificial Intelligence based tools for diagnosis and prognosis of plant components. Colace et al [2] discuss intelligent maintenance tools that improve safety of operations in an electric arc furnace, while Fumigalli et al [3] discuss Agile diagnostic tools based on signature analysis of currents and voltages. For the work discussed in this paper, the ALERT diagnostic tool [4] plays a key role. However, it becomes clear that just forming a diagnostic tool is not enough for preventing failures and breakdowns.

Industries are focussing more on reducing cost and improving production efficiency, plant stability and product quality. AI and ML has been widely accepted as the next paradigm for improving production along all these objectives [5]. Many state of art classification algorithms like SVM & Nearest Neighbours have been applied to remove noisy data and provide accurate results [6]. Many such algorithms contribute significantly in failure diagnosis and process prognostics, e.g. Ravi and Bhattacharya have used Neural Networks (NN) and Extreme Learning Machines (ELM) for predicting breakouts in a Continuous Casting Plant in steelmaking [7] and later used Adaptive techniques based on Reinforcement Learning for modulating the parameters of these models to neutralize process drift [8]. Experiments have been performed, such as incorporating ML techniques to an Electric Arc Furnace (EAF) for optimization of parameters in the main process, successfully lowering energy consumption and increasing productivity [9].

An EAF does not always work in a consistent manner, and there may be some irregularities in its operation. These irregularities lead to the formation of what are called *regimes* of operations for the EAF. A regime indicates a particular level of performance values and condition of the EAF. The question that comes up in this regard is, what are the implications / interpretations of these regimes in terms of the above-mentioned production objectives, and how do we find them? When the EAF operating in a particular regime abruptly shifts to another, it can be understood that something within the system is causing a sort of fluctuation which can then be looked into by the plant operations engineers. As for finding the regimes, there could be alternative approaches. The clustering algorithms, by their nature, are able to separate a given dataset into smaller subsets with similar characteristics [10]. Using these, the regimes can be differentiated with distinct properties.

Two well-known algorithms, Fuzzy C-Means [11-13] and DBSCAN [14] have been used widely for many applications in industry. It has been seen that Fuzzy C-Means algorithms have a lot of utility in various industrial applications. It is able to provide membership function indicative of the degree of belonging for a data sample (tapped from the operating process) to various clusters. However, for assurance and consistency the authors have also used the DBSCAN algorithm on the same data-sets [15]. Experiments with these two algorithms have also been done in a different field for another classification problem, with similar objectives in mind [16].

In Sec. II, we provide descriptions for the main components from the plant that is involved with this research work. Following this, we discuss the clustering technique employed and the reasoning behind it in Sec. III. Sec IV goes in depth with the implementation of the experiments and the key observations made and noted, improvements made as well as the final results obtained. Sec. V provides the concluding remarks for the results that have been obtained in the previous section.

II. OPERATING MACHINERY AND REAL-TIME DATA ANALYTICS

A. *Electric Arc Furnace*

The Electric Arc Furnace (EAF) is a crucial operational facility of a steel plant – an important link in the steelmaking process chain [17]. By passing high-energy currents between two electrodes through a scrap-metal matrix, this apparatus heats the metal to extreme temperatures and then melts it, where, gradually, high quality metal is tapped from the bottom while slag and related impurities float on top. The slag is taken out by tilting the furnace and the higher grade liquid metal is retrieved subsequently, see fig. 1.

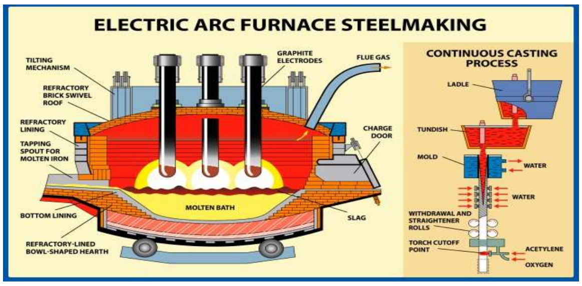

Fig 1. Schematic representation of an Electric Arc Furnace, from multiple sources in open domain.

B. *ID-Fan*

Fumes generated during the steel melting process in the Electric Arc Furnace (EAF) are extracted from the EAF by the suction created by the ID Fan. The suction hood (canopy hood) is also connected to the same ID Fan to collect any secondary fumes to prevent pollution outside the building in which the EAF is located. The ID Fan collects the fumes from the EAF and suction hood and forces it through the Baghouse which consists of filter bags in multiple chambers. The solid particles in the fume are collected in the filter bags, which get dislodged by a reverse air fan and collected at the bottom of the bag chambers. The dust is evacuated from the bottom of each chamber by rotary feeders and screw conveyors. The dust free (clean) fume gets evacuated to the atmosphere through the stack. The process is illustrated in fig. 2.

Analyzing the working condition of the ID-fan gives a direct understanding of the working condition of the EAF of which it is a crucial component. Thus, differentiating the regimes of the ID-fan will in turn enable us to differentiate those of the EAF.

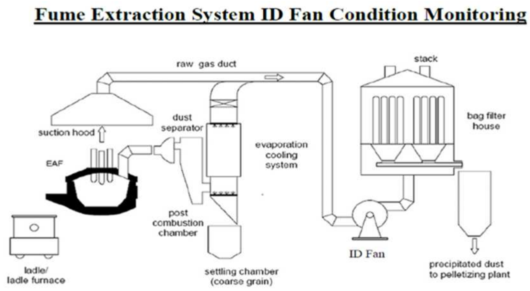

Fig 2. Illustration of an Induced Draught (ID) fan.

C. *ALERT*

Adaptable Leading Edge Reliability Techniques (ALERT), a real time analysis software tool [4], provides an affordable and adaptable platform to develop and customize the most appropriate diagnostic rule applicable to detect anomalous conditions in functioning machinery. It is a state-of-the-art diagnostic tool. Applying knowledge, experience, and historical insights, abnormal trend signatures can be correlated to machinery problems. The architecture of ALERT is essentially a cloud- based application that permits features requiring a large amount of memory space. The functionalities of ALERT are based on real-time data acquired from the plant. A Gateway computer called the ALERT Gateway is installed in the plant to collect relevant analog and digital signals required for the purpose of diagnostics. The signals are sourced from the existing Programmable Logic Controllers (PLCs) or other data sources available in the plant. Fig. 3 provides a high level view of ALERT.

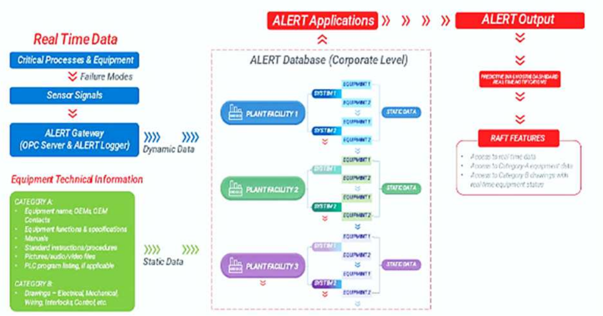

Fig 3. High level view of the ALERT system.

The Gateway acquires the real-time data from the plant-based data sources (PLCs, DCSs, Reflective Memory networks, etc.), validates the data prior to onward transmission to the ALERT cloud via encrypted data packages, utilizing internet or cellular connectivity. The database in the ALERT cloud receives the real-time data, updates itself and shares the data with various Functional Modules. The ALERT system provided sensory data of the ID-fan via its interface for the research. This data consisted of up to 20,000 data points with 6 parameters which are as follows:

1. Motor RMS current of the ID-fan (1 signal)
2. Motor Drive-End and Non-Drive-End bearing vibrations (2 signals)
3. Motor Drive-End and Non-Drive-End bearing temperature (2 signals)
4. EAF tilt angle.

III. CLUSTERING AS A SOLUTION

AI/ML techniques are used worldwide for a wide variety of data analytics and prediction engines. Out of the available techniques, the clustering algorithms appear to be more tailored towards our objective. Clustering helps determine the intrinsic grouping among the unlabeled data present. For instance, it could be interesting to find representatives for homogeneous groups (data reduction), "natural clusters" and describe their unknown properties, suitable groupings or in finding unusual data objects (outlier detection). These algorithms must make some assumptions underlying the similarity of points and each assumption makes different and valid clusters.

A. Fuzzy C-Means Clustering

Fuzzy C-means (FCM) is a data clustering technique in which a data set is grouped into *C* clusters with every data point in the dataset belonging to every cluster to a certain degree. For example, a data point that lies close to the center of a cluster will have a high degree of membership in that cluster, while the same point will have a low degree of belonging to another cluster whose center is far from it. Let

- 'n' be the number of data points
- 'v_j' represent the j^{th} cluster center
- 'm' be the fuzziness index, where $1<=m<=\infty$
- 'c' represent the number of clusters
- 'μ_{ij}' represent the membership of j^{th} data point to i^{th} cluster center, and
- 'd_{ij}' represent the Euclidean distance between j^{th} data and i^{th} cluster center.

The main objective of fuzzy C-means algorithm is to minimize:

$$J(\boldsymbol{Z;U,V}) = \sum_{j=1}^{c}\sum_{k=1}^{N}(\mu_{ik})^m \|z_k - v_i\|_A^2$$

where **Z** is the data matrix with columns representing samples and rows the variables in a sample, **U** is a partition matrix with columns representing samples and rows as the clusters, and **V** is the vector of cluster centroids.

The algorithm works as follows:

1. Initialize a random partition matrix and the number of clusters in a given range.
2. Compute cluster centroids using
$$\mathbf{v}_i^{(l)} = \frac{\sum_{k=1}^{N}\left(\mu_{ik}^{(l-1)}\right)^m \mathbf{z}_k}{\sum_{k=1}^{N}\left(\mu_{ik}^{(l-1)}\right)^m}, 1 \le i \le c$$
here superscript *l* denotes an iterative step
3. Calculate the degree of membership of each data point to each cluster based on their distance from the centroids.
$$D_{ik\mathbf{A}}^2 = \left(\mathbf{z}_k - \mathbf{v}_i^{(l)}\right)^T \mathbf{A}\left(\mathbf{z}_k - \mathbf{v}_i^{(l)}\right)$$
$$1 \le i \le c, 1 \le k \le n$$
$$\mu_{ik}^{(l)} = \frac{1}{\sum_{j=1}^{c}\left(D_{ik\mathbf{A}}/D_{jk\mathbf{A}}\right)^{2/(m-1)}}$$
4. Update the centroids of each cluster by calculating the weighted mean of the data points based on their membership values (as in step 2).
5. Repeat steps 2 and 3 until the membership values converge, or a maximum number of iterations is reached. The convergence condition is given by:
$$\|\mathbf{U}^{(l)} - \mathbf{U}^{(l-1)}\| < \epsilon$$

The degree of membership of each data point is determined by a fuzzifier parameter called *m*, which controls the degree of fuzziness in the clustering. A larger value of *m* leads to more overlapping clusters, while a smaller value of *m* leads to sharper, more distinct clusters.

This clustering algorithm repeatedly runs starting from 1 followed by increments of 1 until the upper limit for number clusters has been reached. Out of all the repetitions, the best optimal cluster number is selected by checking their R_c values, given by:

$$R_c = abs\left(\frac{J_c - J_{c+1}}{J_{c-1} - J_c}\right)$$

$$J_c(\boldsymbol{Z;U,V}) = \sum_{j=1}^{c}\sum_{k=1}^{N}(\mu_{ik})^m \|z_k - v_i\|_A^2$$

B. DBSCAN Clustering

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a popular clustering algorithm that identifies clusters of arbitrary shape in a dataset by grouping together points that are close to each other and have a sufficiently high density of neighboring points. The DBSCAN algorithm does not require specifying the number of clusters in advance but needs to specify a neighbourhood radius and a corresponding minimum density, but it can handle noisy and outlier points. It is widely used in a variety of applications, including image analysis, anomaly detection, and customer segmentation.

The algorithm works as follows:

1. For each point in the dataset, compute its ε-neighborhood, i.e., the set of all points within a distance ε from it.
2. For each point, determine whether it is a core point, a boundary point, or a noise point. A core point is a point that has at least a minimum number of other points (*MinPts*) within its ε-neighborhood. A boundary point is a point that is not a core point but is within the ε-neighborhood of one or more core points. A noise point is a point that is neither a core point nor a boundary point.
3. Create a cluster for each core point and add all of its reachable points (i.e., points within its ε-neighborhood) to the cluster. Repeat this process recursively for all the reachable points until no new points can be added to the cluster.
4. Assign all the boundary points to their respective clusters.
5. Mark all the noise points as outliers.

The key equations governing the DBSCAN algorithm are the ε-neighborhood and the core point definition:

1. ε-neighborhood: The ε-neighborhood of a point p is defined as the set of all points within a distance ε from p.

$$N_\varepsilon(p) = \{q \in D \mid dist(p,q) \leq \varepsilon\}$$

2. Core points: A point p is a core point if it has at least MinPts other points within its ε-neighborhood.

$$|N\varepsilon(p)| \geq MinPts$$

C. Dataset

The ID-fan is a significant component of the EAF and the operating condition of the EAF is closely tied to some of the operational parameters of the ID-fan. For this research work, sensor data of the ID-fan was observed and acquired via the ALERT interface. The data consists of sensor readings picked up throughout the operation of the ID-fan that are crucial to its operation. It had mainly 6 variables:

- Motor *RMS current* of the ID-fan (1 signal)
- Motor Drive End (DE) bearing *vibration*
- Motor Non Drive End (NDE) bearing *vibration*
- Motor Drive End bearing *temperature*
- Motor Non Drive End bearing *temperature*
- EAF tilt angle, while not a reading for the ID-fan, was also taken for greater operational insight.

The tilt angle is indicative of the phase of the EAF. Small tilt angle values, ranging between 0-2 degrees, are observed when the furnace is vertical and in the process of melting. Larger tilt angles indicate that the furnace is tapping out the molten iron or slag.

IV. IMPLEMENTATION AND RESULTS

In the initial stages of our implementation, an hour of data with a frequency of 1 sec was extracted, once per day for a total of seven days from the ALERT interface. This resulted in acquiring an extremely large volume of noisy data. Therefore, it was further down-sized to two-thirds of its initial size. The fuzzy C-means and DBSCAN algorithms were executed with this data as its input. Both the algorithms were able to successfully separate it into two clean clusters.

On analysis of this clustering, it was found that one cluster consisted of the data points where the Motor RMS current values were close to 0 indicating that the system was non-operational. This significantly dwarfed the other different conditions and clumped all of these cases together into one cluster, and the rest to the other (i.e. non-operational) cluster.

The number of data points in the clusters are as follow: -

- Cluster 0: 1677
- Cluster 1: 240 (non-operational regime)

This led to the formation of a major cluster that contained only data samples of non-active conditions which dominated all the other conditions (both the desirable and undesirable ones). This result successfully confirmed the effectiveness of the clustering algorithms since it was able to differentiate the operational and non-operational regimes of the EAF. Also that both Fuzzy C-Means and DBSCAN produced near-identical results. But at the same time, this brought to attention that for finer distinctions based on different features, the data should not contain any points at non-operational status.

The next set of data that was used revolved around the time frames when the system was suspected to be having irregularities during its operation in the past (June and July 2022). This was carried out by close discussions with some of the operational engineers responsible for the system.

Data spanning an hour around each of the time points where irregularities were suspected to be present, was extracted. When data was collected, the regime for each corresponding data point was not known. This resulted in a dataset consisting of nearly 18,000 points of both regular and abnormal condition samples.

On feeding this input to the Fuzzy C-Means clustering algorithm, 2 clusters were found to be the optimal number of clusters for the algorithm with an upper limit of 10 clusters and based on the R_c value.

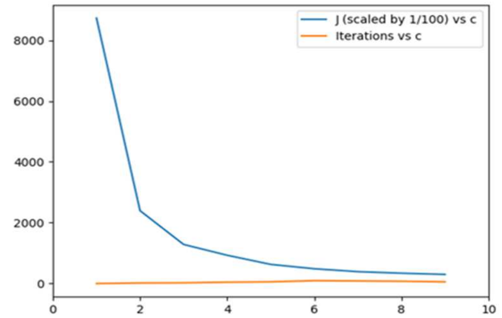

Fig 4. Variation of Objective function value and convergence iterations against number of clusters, for Fuzzy C-Means clustering

For further analysis, 3 clusters were taken into consideration as well. The Fuzzy C-Means algorithm provided centroid values for each cluster, based on which the data samples were classified. The resulting number of data points for each cluster were as follows:

- For 2 clusters separation:
  - C0: 15566
  - C1: 2860

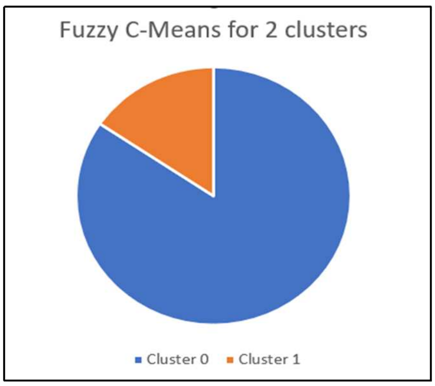

Fig 5. Cluster sizes when only 2 clusters are created.

- For 3 clusters separation:
  - C0: 2860
  - C1: 9115
  - C2: 6451

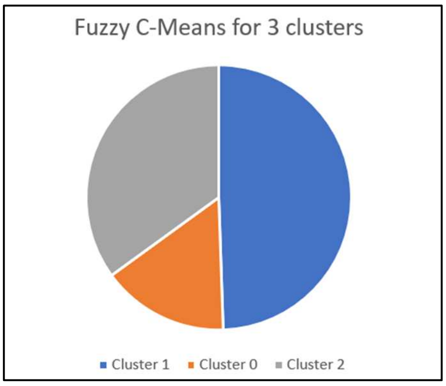

Fig 6. Cluster sizes when 3 clusters are created.

T-distributed Stochastic Neighbor Embedding (t-SNE) is a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions. Specifically, it maps each high-dimensional object into a two- or three-dimensional point in such a way that similar objects are mapped into nearby points and dissimilar objects are mapped to distant points, with high probability. The t-SNE Algorithm was employed to aid in visualizing the 5-parameter clustered dataset so as to draw out unseen relations between the clusters. This is shown in fig. 7 below and it can be seen that all three clusters are neatly separated in their 2-D maps.

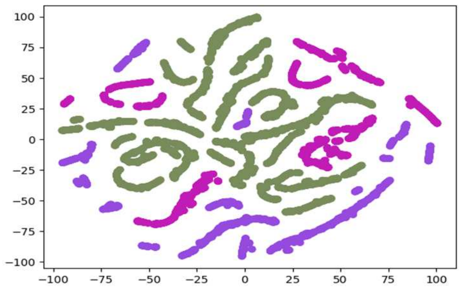

Fig 7. Mapping 5D input variable space into 2D, post clustering, using t-SNE algorithm.

The key takeaway from these results is that *one of the clusters from both the 2 cluster classification and 3 cluster classification, was absolutely identical*. They contained exactly the same data points.

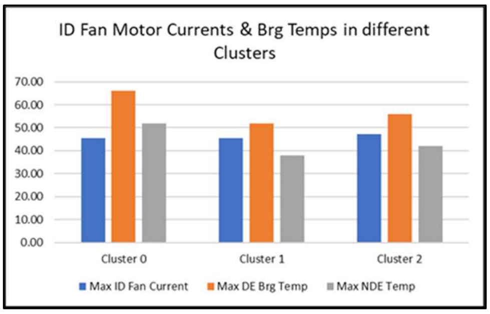

Fig 8a. Max values of ID Fan Motor RMS currents, and bearing temperatures, at each of the three synthesized clusters. DE and NDE denote Drive-End, and Non-Drive-End. Brg denotes Bearing.

On analysis of the 3-cluster classification, several observations were noted. For each cluster the maximum, average and minimum of the bearing temperatures and bearing vibrations were calculated. On comparing the

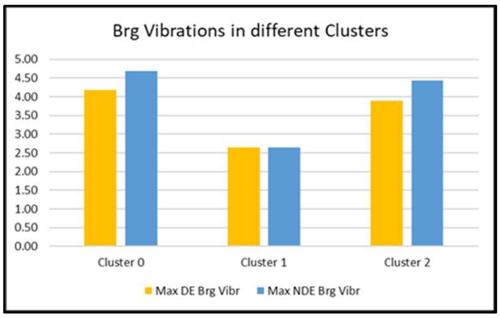

Fig 8b. Max values of Bearing vibrations, at each of the three synthesized clusters.

maximum, average and minimum values of all three clusters, no significant difference was found. The unfiltered data includes bearing temperature and vibration data when the EAF is non-operational (particularly when the EAF is being tapped). So, the data was filtered for each cluster to select data corresponding to tilt angle 0 - 2 degrees, which resulted in getting the data when the EAF is vertical and operational. The filtered comparison gives us a much better contextualization of the data. Bar graphs with the max values were created because that is what appears to be most relevant, and are shown in figs. 8a and 8b.

The filtered comparison tells us the following:

- The max values of the motor currents are very similar in all the three clusters.
- Max bearing vibrations and bearing temperatures are higher in cluster 0 and cluster 2. Cluster 0 shows the higher values between the two.

The data was compared, day by day for the entire month of June and it was found that the bearing vibrations were rising to above 4 mm/sec till June 16th. After that for the months of June and July the vibrations are within normal range with no data above 2.5 mm/sec.

It becomes clear that since Cluster 0 data falls within a data range when vibrations and temperatures are going high, the max values in cluster 0 are the highest. Cluster 1 data falls in the data range when the parameters are normal. Cluster 2 data consists of one single day which had high vibration and temperature, and the rest from the normal period. This explains why the Cluster 0 data presents the highest max values, followed by cluster 2.

For the 3-clusters classification, it was found that the average degree of belonging of samples in Cluster 1, to the Cluster 0, is 0.0121, while the average degree of belonging of samples in Cluster 2 to the Cluster 0, is 0.0527. Implying that samples in Cluster 2 have got higher affinity towards Cluster 0; as compared to affinity of samples in Cluster 1 to Cluster 0. This information can be gleaned from the natural characteristics of the Fuzzy C-Means algorithm.

The clustering brought to our attention some abnormality occurring in some of the days. As far as the process dust and fume load is concerned, it was normal as seen from the motor currents in all the three clusters. Therefore, there has to be something happening in the drive system – implying the clustering mechanism provides inputs to plant operations and maintenance engineers to focus their analysis on specific equipment and sub-processes for drifting behavior.

It is noteworthy that all the above results were obtained from the Fuzzy C-means clustering algorithm alone. The DBSCAN algorithm completely failed to generate any kind of physically meaningful clusters from this noisy industrial data. The authors spent much more time (compared to Fuzzy C-Means) on different versions of DBSCAN, but were unable to produce the desired distinctive clusters.

It may be recalled that for the initial dataset when the Fuzzy C-Means algorithm formed clusters between operational and non-operational conditions, DBSCAN was also able to form precisely the same clusters to separate the two regimes. In this case the distinction was gross. When it came to finer distinctions between different features with significant intrinsic noise, as required from the second dataset, DBSCAN failed while the Fuzzy C-Means worked. This is the second most significant take home from this study. The first, of course, is achieving the desired distinction between abnormal and normal working regimes, which is of great value to the concerned industry.

V. Conclusions

It is demonstrated that Fuzzy C-means Algorithm works well in identifying different regimes of operation of an Electric Arc Furnace, based on fine distinctions extracted from noisy real time industrial data. The DBSCAN algorithm can synthesize realistic clusters under conditions of gross variation between data characterizing the different regimes, but surprisingly, fails in cases of finer distinctions. Furthermore, Fuzzy C-Means provides more information and insight in terms of identifying the degree of belonging of each streaming data sample (extracted from real time operations) to different operating regimes and hence can point towards impending regime transitions – which is of immense value to operations.

Acknowledgment

Authors acknowledge the kind support provided by Mr. Sushil Baraskar of EVSL Steel Plant, Wardha, India.

References

[1] Mohd Javaid, Abid Haleem, Ravi Pratap Singh, and Rajiv Suman, "Artificial Intelligence applications for industry 4.0: A literature-based study", J. of Industrial Integration and Management, Vol. 7, No. 1, 2022, pp. 83-111.

[2] Colace C., Fumagalli L., Pala S., et al. (2013). An intelligent maintenance system to improve safety of operations of an electric furnace in the steel making industry. Chem Eng, Vol. 33, pp 397–402.

[3] Fumagalli L., Ierace S., Dovere E., Macchi M., Cavalieri S., Garetti M. (2011). Agile diagnostic tool based on electrical signature analysis, IFAC Proceedings Volumes (IFAC-PapersOnline), 18 (PART 1), pp. 14067-14072.

[4] https://www.mcartech.com/wp-content/uploads/2022/02/MCARTech-Tech-Deck-for-Site.pdf ; last access March 29, 2023.

[5] Bhattacharya, A.K., Agarwal, A., Nag, S., Hazra, S., Amit, R.K., *Method and System for obtaining perpetual optimum performance of industrial reactor using Game Theory principle*, Indian Patent, Published 2021, App. No. 202131029266.

[6] Limei Liu, Anna Wang, Mo Sha, Xiyan Sun, "Optional SVM for Fault Diagnosis of Blast Furnace with Imbalanced Data", ISIJ International, Sep 2011, Vol. 51, No. 9, pp. 1474-1479, DOI:10.2355/isijinternational.51.1474 .

[7] R. Ravinithesh, A.K. Bhattacharya and G. Rishita, "Advance Predictions of critical digressions in a noisy industrial process-performance of Extreme Learning Machines versus Artificial Neural Networks", IFAC-PapersOnLine, Volume 51, Issue 1, 2018, Pages 98-105, DOI: https://doi.org/10.1016/j.ifacol.2018.05.017

[8] R. R. Annapureddy, A. K. Bhattacharya and Niranjan Reddy, "Adaptive Critic Design for Extreme Learning Machines applied to noisy and drifting industrial processes," 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 2018, pp. 327-334, 10.1109/SSCI.2018.8628664

[9] Vaso Manojlović, Željko Kamberović, Marija Korać, Milan Dotlić, "Machine learning analysis of electric arc furnace process for the evaluation of energy efficiency parameters", Applied Energy, Vol. 307, Feb 2022, DOI: 10.1016/j.apenergy.2021.118209

[10] Estela Ruiz, et al, "Machine Learning Methods for the Prediction of the Inclusion Content of Clean Steel Fabricated by Electric Arc Furnace and Rolling", Metals, Vol. 11, No. 6, DOI: 10.3390/met11060914

[11] Dunn, J. C., "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters". Journal of Cybernetics. 1973, **3** (3): 32–57. DOI :10.1080/01969727308546046

[12] Bezdek, James C., *Pattern Recognition with Fuzzy Objective Function Algorithms,* 1981, ISBN 0-306-40671-3.

[13] R.Suganya and R.Shanthi, "Fuzzy C- Means Algorithm- A Review", J. of Scientific and Research Publications, Vol.2, No. 11, Nov. 2012, ISSN 2250-3153.

[14] https://en.wikipedia.org/wiki/DBSCAN ; DBSCAN Algorithm.

[15] Boonpipob Napasiripakorn, Wilaiporn Lee and Kanabadee Srisomboon, "Investigation of DBSCAN Data Clustering For MCA Cooperative Spectrum Sensing", *2022 International Conference on Power, Energy and Innovations (ICPEI)*, 2022, pp. 1-4, DOI: 10.1109/ICPEI55293.2022.9987052.

[16] Kwang Baek Kim , Doo Heon Song and Hyun Jun Park, "Intelligent Automatic Segmentation of Wrist Ganglion Cysts Using DBSCAN and Fuzzy C-Means", Diagnostics, Dec 2021, DOI: 10.3390/diagnostics11122329.

[17] https://en.wikipedia.org/wiki/Electric_arc_furnace , Electric Arc Furnace