Ensemble of extreme learning machines for diagnosing bearing defects in non-stationary environments under class imbalance condition

Roozbeh Razavi-Far, Mehrdad Saif
Department of Electrical and Computer Engineering
University of Windsor
401 Sunset Avenue, Windsor, N9B 3P4, Canada
Email: {roozbeh, msaif}@uwindsor.ca

***Abstract*—Two practical inevitabilities for diagnostic systems are the abilities of incremental learning in non-stationary environments and diagnosing under the class imbalance condition. The class imbalance condition has been widely occurred in real applications where system usually works in the normal state and it is not easy to collect the representative patterns of faulty classes. This work aims to adapt two state-of-the-art ensemble-based techniques for incremental learning and diagnosing faults in non-stationary environments under the class imbalance. These techniques train several extreme learning machines to create the ensemble which can incrementally learn the relation between features and faults in various class-imbalanced chunks of data collected from non-stationary environments. These diagnostic schemes are applied to diagnose bearing defects in induction motors.**

I. Introduction

With the constant growth in complexity of industrial systems, reliable and safe operations of diagnostic systems are of paramount importance for the environment and human health management, and for the economic process. Besides, the proper performance of the industrial systems has a deep effect on the total cost and quality of the products [1]. A prompt detection of anomalies and diagnosis of faults can avoid performance degradation, system breakdown, and damage to the machinery or human life [2].

Contemporary diagnostic systems usually make use of process models and require a model of the normal process status to detect an anomaly [1]. However, attaining precise and representative models for complex operational processes is a very challenging task [2]. To overcome this problem, a great number of data-driven diagnostic schemes have been designed in the past decade, based on computational intelligence and machine learning approaches [3]–[5].

The data-driven diagnostic schemes are mostly based on intelligent classification algorithms where the classifiers are trained, in a supervised fashion, by means of a subset of patterns representing the normal and faulty states and, then, used to predict the future states once unseen sets of data become available. In non-stationary environments, the data are continuously observed and collected in course of time and, then, fed to the diagnostic classifier chunk by chunk or one by one. The subsequent chunk of data may represent a concept drift [6] or even unseen classes of faults [7]–[9] that were not exist in preceding chunks of data. Fault classifiers then require incremental adjustments in course of time [6]–[9].

The former issue, concept drift, can be tackled by resorting to incremental learning algorithms in non-stationary environments such as just in time (JIT) classifier [10] and the ensemble-based classifiers including streaming ensemble algorithm (SEA) [11], dynamic weighted majority (DWM) [12], Learn^{++}.NSE (Non-Stationary Environment) [13]. The recent techniques for learning in non-stationary environments have been carefully studied in [14].

Moreover, the data collection procedure, in practice, usually creates underrepresented subsets with severe class distribution skews where gathering faulty patterns is a very hard and expensive task [15]. This class imbalance issue further complicates the procedures of incremental learning and consequently fault diagnosis in non-stationary environments [16], [17].

In this respect, a state-of-the-art technique, so-called Learn^{++}.CDS (Concept Drift with SMOTE), has been proposed to overcome class imbalanced issue in non-stationary environments [18]. Learn^{++}.CDS combines the Lean^{++}.NSE for incremental learning in non-stationary environments (i.e., the concept drift) with the Synthetic Minority class Oversampling TEchnique (SMOTE) algorithm for learning under class imbalanced condition [18]. Learn^{++}.CDS initially makes use of SMOTE to reduce the imbalance ratio by creating new synthetic samples of the minor class and, then uses Learn^{++}.NSE for learning the drift concept from the re-balanced set of samples [18].

Another state-of-the-art technique to handle class imbalanced data during the incremental in non-stationary environments is Learn^{++}.NIE [17]. Learn^{++}.NIE is an extension of the Lean^{++}.NSE for incremental learning in non-stationary environments under class imbalanced condition. Both of the algorithms belong to a Learn^{++} class of algorithm which is basically an ensemble learning scheme in which an additional set of classifiers is trained with the newly emerging chunk of data and, then, added to the ensemble scheme [17].

These ensemble-based techniques usually outperform single classifiers approaches, particularly with gradual drifts [19]. Therefore, in this work, these two ensemble-based schemes are adopted and compared for incremental learning and diag-

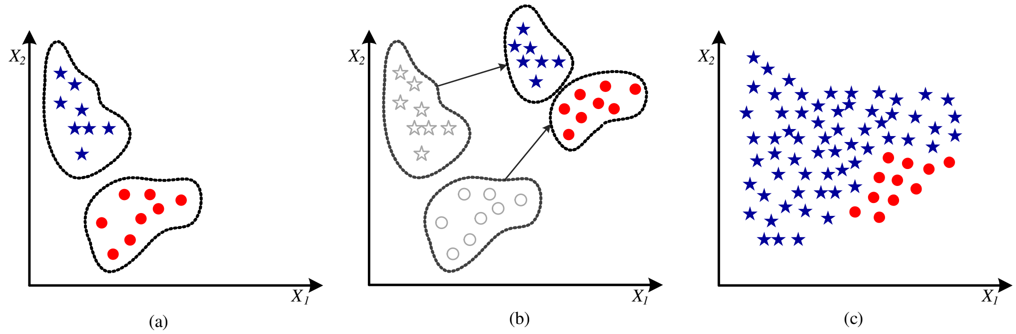

Fig. 1. Illustrative two-class examples: (a) and (b) the concept drift, and (c) the class imbalance problem.

nosing faults in non-stationary environments under the realistic situation of the class imbalanced condition.

These approaches are used to incrementally learn and diagnose the bearing defects in induction motors, where several chunks of imbalanced data gradually collected and become available in course of time resembling a realistic situation of fault diagnosis in non-stationary environments.

The remainder of the paper is structured as follows. The case study and problem of diagnosing bearing defects in non-stationary and class imbalance conditions are briefly explained in Section II. Section III formally presents the the diagnostic schemes; and briefly describes two selected state-of-the-art techniques for incremental learning in non-stationary environments under class imbalance condition as core components of the two diagnostic schemes. These diagnostic schemes are compared together in diagnosing the bearing defects in induction motors and their experimental results are presented in Section IV. Finally, Section V contains conclusions.

II. Problem Definition

The ultimate goal of this work is to design a data-driven diagnostic system which can work properly in non-stationary and class imbalance conditions. A typical data-driven diagnostic scheme usually receive an input vector from a set of observations and predicts a target label for that input representing the predicted state of the system, i.e., faulty or normal state.

The process of designing the diagnostic classifier is basically a supervised learning task which contains two major steps. In the first step, so-called training, fault classifier receives a training subset $T = \{(x_1^T, y_1^T), (x_2^T, y_2^T), \ldots, (x_{m_T}^T, y_{m_T}^T)\}$, where m_T stands for the number of samples in T, and $y_i \in Y \subset \{c_1, c_2, \ldots, c_{w_T}\}$, in which w_T stands for the number of class label in T, and then, trains a classifier model, h representing the relations between inputs and target labels.

In the second step, so-called prediction, it receives a subset of unlabeled data $\Gamma = \{x_1^\Gamma, x_2^\Gamma, \ldots, x_{m_\Gamma}^\Gamma\}$, where m_Γ stands for the number of samples in Γ and y_i^Γ are not known. The previously trained classifier then predicts the target $\hat{y}_i^\Gamma$.

The fault classification is based on the hypothesis that all samples in T and Γ are partially similar, i.e., x_i^T and x_i^Γ are sampled through the same probability distribution. This hypothesis is not valid in non-stationary environments where the samples of several training T_t and test Γ_t subsets are collected in course of time representing different operation conditions, and hence, the statistical characteristics of data in different T_t and Γ_t can change in course of time, i.e., class concept tends to drift in non-stationary environment [17]. In other words, the joint distributions vary in course of time in a way that $p_T(x, c) \neq p_\Gamma(x, c)$.

The performance of the previously trained classifiers, h_t, decrease w.r.t. Γ_{t+1}, as a result of the concept drift. Panels (a) and (b) in Figure 1 illustrates the concept drift for a two-class problem.

Another issue of concern is presence of the severe class distribution skews in T_t and Γ_t which significantly reduce the fault classification performance. This is a very realistic situation in many cases where majority of the collected samples belong to the normal state and only few samples representing a faulty situation. For instance, consider collecting the data from a system process which mostly works under normal condition, while the samples of a faulty class are very rare or very difficult to obtain.

Given a dataset T_t with m_t samples (i.e., $|T_t| = m_t$) with two classes $\omega_T = 2$ (e.g., normal and a faulty state), two majority and minority subsets can be defined as follows $T^{maj} \subset T$ and $T^{min} \subset T$, so that $T^{maj} \cap T^{min} = \{\oslash\}$, $T^{maj} \cup T^{min} = \{T\}$, and $|T^{maj}| = m^{maj} \gg |T^{min}| = m^{min}$, where m^{maj} and m^{min} stand for the number of samples in the majority and minority subsets, respectively. Panels (c) in Figure 1 illustrates the class imbalance condition for a two-class dataset.

To overtake these two issues, two state-of-the-art techniques are considered and adopted for fault diagnosis in non-stationary and class imbalance conditions. These diagnostic schemes are used to diagnose bearing defects in induction motors. The bearing datasets have been taken from the Case Western Reserve University bearing data center [20].

Bearing faults are one of the main causes for machine downtime in industrial process, and thus, various diagnostic schemes have been applied to detect anomalies and diagnose defects in induction motors [21], [22].

This work focuses on diagnosing bearing defects in dynamic situation where different batches of data are collected and become available in course of time. These batches of data contains samples of normal and faulty classes representing different operating conditions, i.e., the concept drift, where these batches of data contain representative samples of different speeds (1797-1720 rpm) for motor loads of (0-3) horsepowers.

Besides, these batches of the data contain samples of faulty classes that were seeded on the rolling elements and on the outer race, where these faults ranging in diameter from 0.007 to 0.028 in (0.18-0.71 mm) resembling the concept drift.

Moreover, the ratio of the faulty and normal samples varies from one dataset to another, representing different class imbalance levels such as highly class imbalanced (HCI) data which resembles presence of anomaly, moderate class imbalanced (MCI) data, and full class balanced (FCB) data which resemble fault progression. Figure 2 illustrates the imbalanced ratio in three differnet datasets representing HCI, MCI and FCM, respectively.

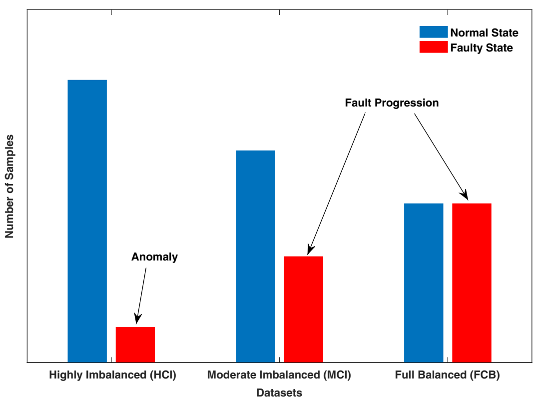

Fig. 2. (HCI) dataset contains few faulty samples resembling the presence of anomaly, while (MCI) and (FCB) datasets contain more faulty samples resembling fault progression.

The detailed description of diagnostic schemes and the experimental setup are given in the following sections.

III. Design of the diagnosis system

The most critical components of the induction motors which cause the majority of failures are the roller bearings, the stator and the rotor, respectively [23]. This has been led to designing numerous vibration-based diagnostic systems which aim to detect and isolate the roller bearing defects [21].

Efficient diagnosis of the bearing defects is mainly based on the proper evaluation of the vibration signal. In data-driven diagnostic frameworks, this has been done by extracting informative and discriminant features from the vibration signal [24]. Thus, the proposed diagnostic system in this work has a signal processing module prior to the fault classification module. Figure 3 depicts the general diagram of the proposed diagnostic system.

A. Signal Processing

The proposed diagnostic system contains a SVD-based signal processing module. This is indeed the Hankel matrix-based SVD [25], [26] which aims to represents a one-dimensional vibration signal in terms of a set of variations independent principal components including possible white noise or different trends.

These extracted principal components can be used a set of features, however, only the most informative principal components are selected as a proper features and the rest are discarded to reduce the dimensionality of the feature set. It is important to note only the first few principal components are informative and explain the most variability of the data and the rest of the components fit noise and cause overfitting.

This procedure has three major steps. It initially embeds the Hankel matrix, arranging the vibration signal of size m into n lagged vector of size m. It then performs singular value decomposition on the obtained matrix of size $m \times n$. It then selects the most important singular values forming a set of smaller size of features in terms of features (the reader can refer to [25], [26], for a more detailed explanation on the signal processing module). Figure 3 shows the different steps of the signal processing module.

The process of the signal processing ends with segmentation of the extracted features and calculating entropy of each segment, reducing the number of samples for the ease of the fault classification task [23].

Here, two eigenvalues (i.e., principal components) are selected since the first two principal components have larger values and they are more distinguishable for different classes of defects, providing a proper set of discriminant features.

B. Dynamic Feature Collection

In this work, two scenarios have been considered resembling the drift concepts in dynamic environments, under class imbalance condition. Each scenario contains three vibration signals $\{S_1, S_2, S_3\}$. These signals are passed through the signal processing module and, then, the extracted features $\{D_1, D_2, D_3\}$ are fed to the fault classification module.

Each feature set includes some samples of the normal state and a specific defect. The structure of these features sets are fully reported in Table I. The sequence of these feature sets in each scenario resembles a non-stationary environment with the drift concept under class imbalance distribution which is a very common situation in practice.

Collected samples in the feature sets $\{D_1, D_2, D_3\}$ of the first scenario resembles a gradual concept drift where each chunk o feature introduces new set of samples representative of the same class but in different motor load and shaft speed (slight change) while the imbalance ratio also varies from HCI to MCI and from MCI to FCB, respectively (See Table I).

On the other hand, collected sample sin the feature sets $\{D_1, D_2, D_3\}$ of the second scenario resemble an abrupt concept drift, where each chunk of feature introduces an abrupt change in both operational condition (i.e., moto load and shaft speed) and width of the defect. The former resembles the sudden drift where the samples vary from the smallest motor load ($0hp$) and the maximum shaft speed ($1797rpm$) to the largest motor load ($3hp$) and the minimum shaft speed ($1730rpm$), while the latter resembles the sudden change in the width of the ball defect (See Table I).

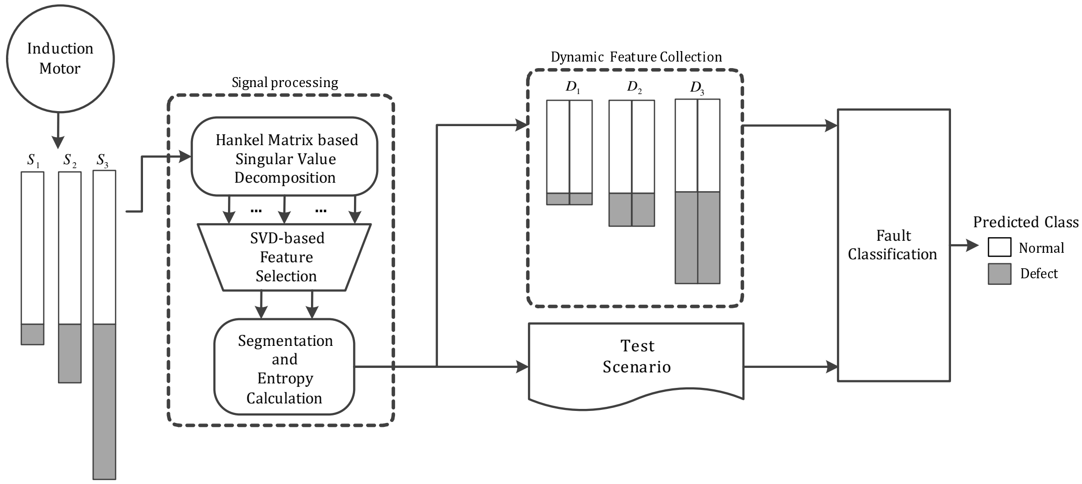

Fig. 3. The designed diagnostic scheme including (1) the signal processing module, (2) dynamic feature collection, and (3) fault classification module.

TABLE I. NUMBER OF SAMPLES IN EACH FEATURE SET IN EACH SCENARIO, FOR EACH DEFECT, WIDTH, LOAD AND SPEED.

Class	Motor Load (hp)	Shaft Speed (rpm)	Defect Width ($in.$)	Scenario 1			Scenario 2		
				D_1	D_2	D_3	D_1	D_2	D_3
Normal	0	1797	-	200	-	-	200	-	-
Normal	1	1772	-	-	200	-	-	-	-
Normal	2	1750	-	-	-	200	-	-	-
Normal	3	1730	-	-	-	-	-	200	200
Outer Race	0	1797	0.007	20	-	-	-	-	-
Outer Race	1	1772	0.007	-	100	-	-	-	-
Outer Race	2	1750	0.007	-	-	200	-	-	-
Ball	0	1797	0.007	-	-	-	20	-	-
Ball	3	1730	0.028	-	-	-	-	100	200
Total	-	-	-	220	300	400	220	300	400
Imbalance Ratio	-	-	-	HCI	MCI	FCB	HCI	MCI	FCB

Figure 4 illustrates the two-dimensional features of $\{D_1, D_2, D_3\}$ in the second scenario w.r.t. the class of each sample, where the number of samples in D_1, D_2 and D_3 represents HCI, MCI and FCB, respectively; and the sequential emergence of these three feature sets resembles an abrupt concept drift. These feature sets are then fed to the fault classification module to be used during the training.

C. Fault Classification

The collected feature sets dynamically emerge as sets of inputs for the fault classification module. In this situation, there is a need for the fault classification module to (1) incrementally learn the novel information to handle the drift concept without retrieving the previously acquired knowledge, (2) preserve previously trained knowledge, (3) construct a set of fault classifiers that can handle class imbalance, and (4) perform learning in the non-stationary environments by resorting to a class-specific (i.e., normal and defect) weighted error instead of using classification error. The latter is due to the fact that the use of classification error can lead to a very low total error by correct classification of the samples of the normal class (i.e., major class) and misclassifying the representative samples of defect (i.e., minor class), where the performance on the minor class (i.e., defect) is of paramount importance to diagnose bearing defect.

1) Learn^{++}.NIE: To address the above issues, here, an algorithm, called Learn++.NIE, is adopted in the fault classification module and applied for diagnosing the bearing defect. Upon emergence of each chunk of features D_t, where the distribution of D_t and D_{t-1} is not necessarily the same $p_t(x, c) \neq p_{t-1}(x, c)$, Learn^{++}.NIE generates a subset of ensemble of fault classifiers. It makes use of a variation of bagging algorithm to train a subset of ensemble of fault classifiers based on all of the samples of the minor class (i.e., defect) and a randomly drawn samples of the major class in D_t [17]. This aims to reduce the imbalance ratio among the samples for better training of the fault classifiers. On the other hand, constructing a subset of ensemble of these fault classifiers helps to learn whole the samples of the major class and avoids loss of information. This variation of bagging merges the trained fault classifiers through majority voting aggregation scheme to form the subset of ensemble.

Learn^{++}.NIE, then, evaluates all trained subsets of ensemble, that are trained thus far, w.r.t. the most recent data, D_t. It then calculates the predicted classes by means of the fault classifiers in the subset of ensemble and, then, measures the recall for each class and, then, computes a class-specific weighted average error. Learn^{++}.NIE makes use of the class-specific

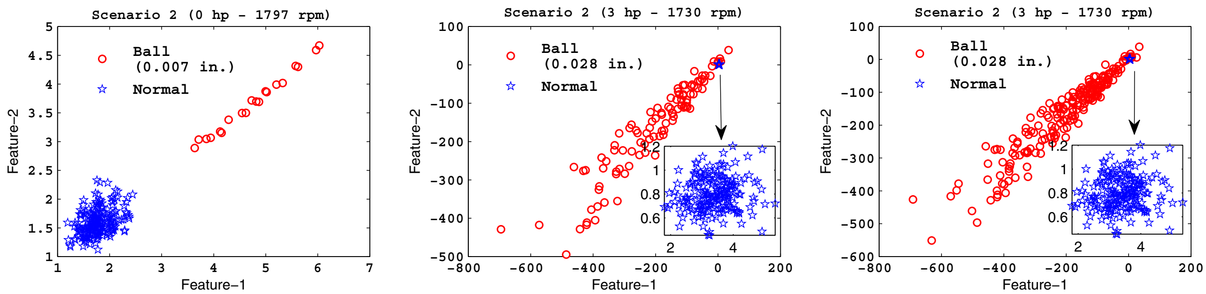

Fig. 4. Presentation of the two-dimensional samples in the second scenario w.r.t. the classes, i.e., normal state and ball defect. The left panel shows the first chunk of samples collected at 0 hp and 1797 rpm (HCI), the central panel shows the second chunk of samples collected at 3 hp and 1730 rpm (MCI), and the right panel presents the third chunk of samples collected at 3 hp and 1730 rpm (FCB).

weighted error to update classifiers weights by increasing the voting weight of a fault classifier with a higher recall value on both of the major and minor classes, i.e., normal and defect states [17].

It finally computes the age-adjusted weighted error average to determine the voting weight of each subset of ensemble and, then, uses the weighted majority of votes of the subsets of ensemble to predict the final class of each sample. The reader can refer to [16], [17], for a more detailed explanation on the Learn^{++}.NIE. Figure 5 illustrates the general scheme of the fault classification module, i.e., Learn^{++}.NIE.

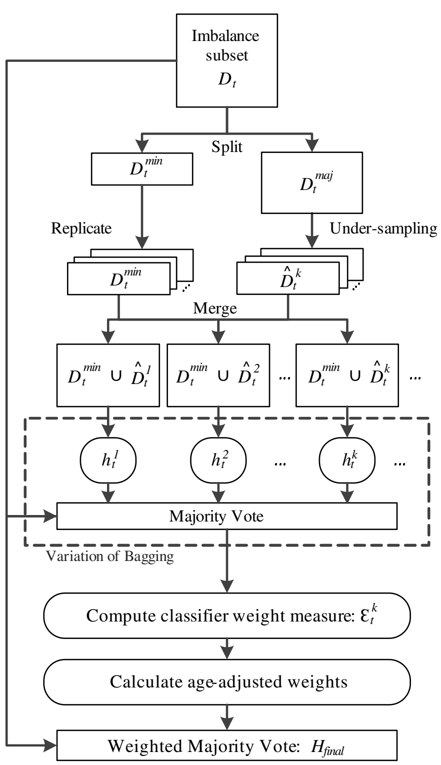

Fig. 5. Block diagram of the Learn^{++}.NIE algorithm

The extreme learning machine (ELM) has been used as individual fault classifier to form the subset of ensemble. The ELM-based fault classifier is a single hidden layer feedforward network [27]. ELM can merge the activation functions within the hidden nodes and randomly initializes the parameters of the hidden neurons regardless of the pairs of inputs. It analytically updates the output weights by optimal combination of the incoming signals from the hidden layer [27].

2) Learn^{++}.CDS: The fault classification module can also be constructed by resorting to Learn^{++}.CDS instead of Learn^{++}.NIE. Learn^{++}.CDS receives a chunk of features (D_t) at each time stamp. It assigns a uniform weight to each samples of the first chunk of features, otherwise (i.e., for the upcoming chunk of features) it evaluates the newly emerged feature set with the existing ensemble of classifiers and, then, updates and re-normalizes the weights of samples by means of the classification error. It consequently increases the weights of the misclassified samples. It then makes use of SMOTE to generate new synthetic samples of the minor class (i.e., defect) and creates a re-balanced feature subset, and consequently, makes use of the newly re-balanced subset to train a fault classifier and adds it to the ensemble. It then evaluates all the trained fault classifiers w.r.t. D_t and calculates their error. The rest of the Learn^{++}.CDS algorithm is very similar to the Learn^{++}.NIE algorithm to obtain the predicted class.

Learn^{++}.CDS does not follow a sample selection strategy, but rather use of the weights of samples to control penalty distribution to compute voting weights. It also makes use of the dynamic (time adjusted) voting weights to reinforce the role of the classifiers that outperform other ensemble members on the most recent data even supposing these classifiers have been trained in the previous cycles.

IV. Diagnostic Results

The results of these two algorithms are compared w.r.t. classification of the bearing defects in the proposed scenarios. Each subset of ensemble contains five extreme learning machines. Two cases have been considered for training the Learn^{++}.NIE algorithm by adjusting its recall parameter: (a) same penalty value has been given to both the normal state and defect, (b) more penalty has been given to the defect class. Figure 6 illustrates recall of both algorithms w.r.t. to various subsets of each scenario.

The attained recall results show perfect performance of both Learn^{++}.NIE and Learn^{++}.CDS on the feature subsets of the first scenario, regardless of parameter tuning in the former algorithm. This is due to the distribution of samples of the normal state and the outer race defect in the subsets of the first scenario, where the samples of the outer race defect are easily

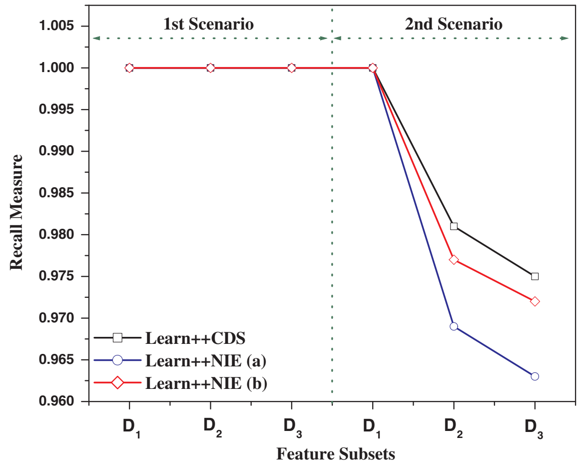

Fig. 6. Recall measure comparison on both scenarios.

separable from the samples of the normal state. The outer race is the first location which is subject to spalling defect. Both techniques have a perfect recall measure with respect to classification of samples of D_1 in the second scenario, however, their recall measures decrease in classifying samples of D_2 and D_3 in the second scenario, which is due to abrupt change in the motor load and width of the ball defect. Although both techniques experience a slight reduction in recall measures, their performances are still satisfactory and comparable. Figure 6 also show that assigning a larger penalty to the ball defect improves the recall measure. In general, Learn^{++}.CDS slightly outperforms Learn^{++}.NIE, however, it has a larger computational time to generate synthetic samples of the minor class to re-balance the feature subsets and also it needs user knowledge about tuning the parameter of SMOTE which is not always available in the online monitoring applications.

V. Conclusion

This work aims to develop a multi-level scheme to diagnose bearing defects in induction motors. It includes three modules: signal processing, dynamic feature collection and fault classification. Signal processing is a Hankel matrix-based singular value decomposition of the vibration signal. The fault classification module adapts two ensemble-based algorithms that can incrementally learn and diagnose bearing defects in non-stationary environments under the class imbalance.

References

[1] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault Tolerant Control*. Springer-Verlag, Heidelberg, 2003.

[2] R. Razavi-Far, H. Davilu, V. Palade, and C. Lucas, "Model-based fault detection and isolation of a steam generator using neuro-fuzzy networks," *Neurocomputing Journal*, vol. 72, pp. 2939–2951, 2009.

[3] A. Vemuri and M. Polycarpou, "Neural-network-based robust fault diagnosis in robotic systems," *IEEE Transactions on Neural Networks*, vol. 8, no. 6, pp. 1410–1420, 1997.

[4] P. Baraldi, R. Razavi-Far, and E. Zio, "Bagged ensemble of FCM classifier for nuclear transient identification," *Ann. Nucl. Energy*, vol. 38, pp. 1161–1171, 2011.

[5] R. Razavi-Far, H. Davilu, V. Palade, and C. Lucas, "Neuro-fuzzy based fault diagnosis of a steam generator," in *7th IFAC Conf. on Fault Detection, Supervision and Safety of Technical Processes*, Barcelona, Spain, 2009, pp. 1180–1185.

[6] P. Baraldi, R. Razavi-Far, and E. Zio, "Classifier-ensemble incremental-learning procedure for nuclear transient identification at different operational conditions," *Reliab. Eng. Syst. Safe.*, vol. 96, pp. 480–488, 2011.

[7] R. Razavi-Far and M. Kinnaert, "Incremental design of a decision system for residual evaluation: A wind turbine application," in *8th IFAC Conference Fault Detection, Supervision and Safety of Technical Processes*, vol. 8(1), 2012, pp. 343–348.

[8] R. Razavi-Far, P. Baraldi, and E. Zio, "Dynamic weighting ensembles for incremental learning and diagnosing new concept class faults in nuclear power systems," *IEEE Trans. Nuclear Science*, vol. 59, pp. 2520–2530, 2012.

[9] R. Razavi-Far and M. Kinnaert, "A multiple observers and dynamic weighting ensembles scheme for diagnosing new class faults in wind turbines," *Control Engineering Practice*, vol. 21(9), pp. 1165–1177, 2013.

[10] C. Alippi and M. Roveri, "Just-in-time adaptive classifiers - part i: Detecting nonstationary changes," *IEEE Transactions on Neural Networks*, vol. 19, no. 7, pp. 1145–1153, 2008.

[11] W. Street and Y. Kim, "A streaming ensemble algorithm (sea) for large-scale classification," in *Proc. 7th ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining*, 2001, pp. 377–382.

[12] J. Kolter and M. Maloof, "Dynamic weighted majority: An ensemble method for drifting concepts," *J. Mach. Learn. Res.*, vol. 8, pp. 2755–2790, 2007.

[13] R. Elwell and R. Polikar, "Incremental learning of concept drift in nonstationary environments," *IEEE Trans. Neural Netw*, vol. 22, no. 10, pp. 1517–1531, 2011.

[14] G. Ditzler, M. Roveri, C. Alippi, and R. Polikar, "Learning in nonstationary environments: A survey," *IEEE Computational Intelligence Magazine*, vol. 10, no. 4, pp. 12–25, 2015.

[15] R. Gouriveau, E. Ramasso, and N. Zerhouni, "Strategies to face imbalanced and unlabelled data in PHM applications," in *Chemical Engineering Transactions*, vol. 33, 2013, pp. 115–120.

[16] G. Ditzler and R. Polikar, "An ensemble based incremental learning framework for concept drift and class imbalance," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, 2010, pp. 1–8.

[17] ——, "Incremental learning of concept drift from streaming imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 10, pp. 2283–2301, 2013.

[18] G. Ditzler, R. Polikar, and N. Chawla, "An incremental learning algorithm for non-stationary environments and class imbalance," in *Int. Conf. on Pattern Recognition (ICPR)*, 2010, pp. 2997–3000.

[19] T. Hoens, R. Polikar, and N. Chawla, "Learning from streaming data with concept drift and imbalance: an overview," *Prog. Artif. Intell.*, vol. 1, no. 1, pp. 89–101, 2012.

[20] "Case western reserve university bearing data center website." [Online]. Available: http://csegroups.case.edu/bearingdatacenter/home

[21] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the case western reserve university data: A benchmark study," *Mech. Sys. & Sig. Proc.*, vol. 64-65, pp. 100 – 131, 2015.

[22] B. Li, M. Chow, Y. Tipsuwan, and J. Hung, "Neural-network-based motor rolling bearing fault diagnosis," *IEEE Trans. Ind. Elec.*, vol. 47, pp. 1060–1069, 2000.

[23] M. Farajzadeh-Zanjani, R. Razavi-Far, and M. Saif, "Efficient feature extraction of vibration signals for diagnosing bearing defects in induction motors," in *Proceedings of IEEE Joint Congress on Neural Networks (IEEE IJCNN)*, 2016, pp. 4505–4511.

[24] M. Farajzadeh-Zanjani, R. Razavi-Far, M. Saif, J. Zarei, and V. Palade, "Diagnosis of bearing defects in induction motors by fuzzy-neighborhood density-based clustering," in *IEEE Int. Conference on Machine Learning and Applications (ICMLA)*, 2015, pp. 935–940.

[25] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure - SSA & Related Techniques*. Chapman & Hall, 2001.

[26] M. Kang and J.-M. Kim, "Singular value decomposition based feature extraction approaches for classifying faults of induction motors," *Mechanical Systems and Signal Processing*, vol. 41, pp. 348 – 356, 2013.

[27] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, pp. 489–501, 2006.