How to Eliminate Missed Problems and False Alarms Using Machine Learning for Vibration Monitoring and Analysis




Predictive maintenance, and specifically asset condition monitoring and analysis in metal production plants, continues to see significant advances in cloud computing, sensors, and machine learning technologies. Primetals Technologies and ITR have partnered to provide the latest predictive maintenance services and solutions and ongoing research and development in waveform analytics. After decades of using human and statistical techniques, machine learning methods are now being applied to further improve efficiency and accuracy in the diagnosis and prognosis of potential failure modes of rotating machinery. Machine learning methods allow monitoring systems to react even faster and more precisely than traditional tools and methods. This paper discusses how machine learning was successfully applied to route-based data collection and analysis and condition monitoring and analysis systems for machine vibration data to improve process efficiencies and system response times. Additionally, it demonstrates how machine learning further augments even the most experienced professional analysts to ensure no missed problems and false alarms.

Authors

Borui Li (top row, left)
Research Engineer, ITR,
Bethlehem, Pa., USA
lib@itr.com

Andrew Lauden (top row, right)
Chief Technology Officer, ITR,
Bethlehem, Pa., USA
laudena@itr.com

Jonathan Davis (bottom row, left)
Chief Executive Officer, ITR,
Bethlehem, Pa., USA
davisj@itr.com

Klaus Stohl (bottom row, right)
Head of Condition Monitoring,
Metallurgical Services, Primetals
Technologies, Linz, Austria
klaus.stohl@primetals.com

In the 1970s, predictive maintenance (PdM) played a very small role in maintenance practices for most industries with exception, perhaps, to the petrochemical industry. While the technology existed to collect, store and trend important machine criteria such as temperature, pressure, vibration, oil content, etc., it was more commonplace to "run to breakdown" or perform time-based maintenance during extended shutdowns. The 1980s saw the wide acceptance and implementation of predictive maintenance programs, due in large part to advancements in sensors, personal computing and portable monitoring technologies. During the 1990s, sensor technology improved further and the size and cost of sensors decreased. With improved technology and lower costs, dedicated monitoring systems have become cost-effective alternatives to portable monitoring for select applications.

By definition, predictive maintenance is a maintenance practice that involves monitoring one or more machine criteria to assess the condition of that machine. The most commonly used predictive technologies include oil analysis, infrared thermography, airborne ultrasonic detection and vibration analysis. This paper discusses advancements in vibration analysis technology with specific emphasis on how condition monitoring and analysis systems (CMAS) use digitized knowledge to develop effective machine learning algorithms, and by extension, how these systems function as smart monitoring and analysis systems in the steel industry.

Traditional vibration monitoring and analysis has used one or both of the following methods: (a) establishing and monitoring against thresholds related to individual vibration characteristics, such as overall root mean square (RMS), banded RMS, or peak vibration, and (b) full-signature analysis whereby an expert reviews each frequency spectrum, time waveform and key feature, and relates it to a potential failure mode. The former approach is efficient and can direct analysts to potential problems requiring further analysis. However, when the analysis involves complex assets that

This article is available online at AIST.org for 30 days following publication.

experience varying conditions, such as changes in load, speed and product, resulting in inconsistent vibration, this approach quickly fails. This is particularly true for machine faults that exhibit through low amplitude and high-frequency vibration. The later approach, full-signature analysis inclusive of analysis of all feature data, is far more thorough and returns more precise and timely results, but it also consumes more resources.

Advances in cloud computing, precision sensors and machine learning now make it possible to apply both vibration monitoring and analysis approaches on complex steel assets while consuming considerably fewer resources. By digitizing the knowledge of the experienced analyst over a broad set of asset operating conditions and using this knowledge base as the machine learning training sets, new algorithms can provide early detection of developing problems that were only previously possible with full analysis.

The general approach is to use analysis by experts to categorize machine conditions as either healthy or unhealthy based on a broad set of vibration signals, inclusive of feature data, time domains and frequency spectra. Further, if a machine is unhealthy, categorize the potential fault condition by severity and type(s) (e.g., misalignment, imbalance, bearing issue, gear issues). Machines categorized as healthy are given positive labels and machines categorized as having potential faults are given negative labels. The data related to machines with negative labels are added to the training set.

While the general approach is simple in concept, developing sufficient training sets is not easy because the traditional analysis methods taught and practiced in the vibration analysis industry often result in incomplete data sets insufficient for full analysis. Any alarm-based screening methods, as are common and generally accepted as good practice, ultimately result in incomplete or non-analysis and gaps in feature data. Fortunately, ITR developed processes in the 1980s specifically designed for analyzing variable load and speed equipment in the metals processing industry, and this process deviates significantly from normal industry practice. The result is comprehensive feature data and information (results of analysis) sets without gaps.

Even with complete and comprehensive data sets, another obstacle is unbalanced data. As expected, there are always far more healthy machines with positive labels than machines with negative labels and potential failure modes. Moreover, different machines, and even individual measurement locations on the same machine, have their own signatures and signature characteristics. For example, a spectrum on machine A can indicate a healthy condition, but a similar spectrum on machine B would indicate a potential issue. Process and asset parameters have significant impact on the classification of data. Consequently, machine learning models were fit to individual measurement locations on every machine. Sometimes, comprehensive "negative" label data sets did not exist because not all faults had occurred on all machines.

To solve these problems, ITR developed a machine learning exception system based on route-based data collection and analysis (RDCA) (RDCA processes collect data manually at intervals more frequent than typical failure modes and experts analyze the data and report findings). The expert findings were digitized, verified and then labeled accordingly as positive or negative using proprietary ITR software. Using this training method, the machine learning model was able to overcome the problem of unbalanced data sets to develop a universal scheme which was then implemented across most machine types, by each type of measurement location (e.g., drive-end bearing). Using the historical feature data stored in the ITR PdM database on the ITR cloud and inspired by Reference 7, a data set of 57,996 vibration signals was built. Using the signals related to the "positive" labeled data on each measurement location, the data set was searched to find the signals with high correlation to form a qualified training data set for a one-class support vector machine (OSVM) model.[17,18] The training for OSVM only needs positive signals as input. Utilizing median absolute deviation (MAD) and OSVM training method with the data set, a tight decision boundary was formed which led to a very sensitive model. But considering the goal of accurately discerning the large number of good machines from the much smaller set of machines with potential issues, an oversensitive model was acceptable and suitable.

For the negative signals identified by OSVM, a convolutional neural network (CNN) was designed and implemented to analyze the signatures. This included applying an aggregating logic to integrate each measurement location's signatures and then output a diagnosis for the machine as one of the following fault conditions: "Bearing Vibration," "Coupling," "Load or Process," "Looseness" or "Signal Abnormal." With the statistics extracted from many years of historical data, a severity from 1-star (minor), 2-star (moderate) and 3-star (severe) was assigned. Furthermore, a semantics analysis technology was designed and implemented, enabling the models to output the results as a simple report. The testing results in the period from December 2020 to February 2021 for 956 machines showed the target system automatically labeled 728 machines (76.15%) as "OK" with a 99.9% accuracy. Of the machines identified and verified as having issues, the specific diagnosis correctly classified the fault condition with an accuracy of 89.4%.

Additionally, ITR also developed another machine learning algorithm and model for a CMAS. With some CMASs, expert analysis and reporting may not be routine and periodic, as was the case with the installed ITR system used in the study. Instead, experts respond to alarms (exception notifications) and analyze alarm/exception data. The CMAS will raise an alarm if any feature values exceed pre-set thresholds. Feature values can be discrete, calculated, or derived statistically, logically or historically. Common feature values include overall RMS, banded RMS, peak, crest factor, intensity factor, etc. While systems that include routine and emergency analysis provide larger data sets of verified results, CMASs provide much larger vibration data sets. With a CMAS, data is collected and processed every few minutes under normal operating conditions to assess machine condition. Due to the abundance of vibration data, no separate data set was needed in this case.

Just as with the RDCA system, ITR used the machine learning method OSVM and the dynamic trust model (DTM)[2,9,10,13] and used data for the training set from a CMAS deployed in a pickle line tandem cold mill (PLTCM) for the purpose of identifying traditional drivetrain issues as well as fifth-octave chatter issues. This particular study focused on the latter issue. In this case and through expert analysis, specific characteristics of the impulse peaks at known calculated frequencies and amplitudes were related to a specific known potential chatter issue. More specifically, these issues were directly related to early bearing faults and spalls in the rolls of the mill system. The OSVM on the CMAS data sets was able to identify the specific impulse peaks that the traditional signature threshold method could not. During the three months after the machine learning algorithm was deployed on the PLTCM CMAS, one bearing issue was diagnosed from harmonics and four spall issues were diagnosed from impacting. Additionally, there were no false alarms.

Related Work

There are some but not many published research efforts related to using machine learning for vibration analysis; e.g., Gan et al. in Reference 6 are using deep learning methods to recognize bearing faults on a machine which includes a 2-hp motor, a torque transducer, a dynamo-meter and a loaded motor. Same as Reference 6, Luo et al. in Reference 11 utilize deep learning methods such as an auto encoder, artificial neural network (ANN) and CNN to provide early detection on impulse responses from time waveforms from the bearings of rotors. Lepine et al. in Reference 8 also aim to detect the impulse response (i.e., shocks) in vibration signals by using the support vector machine (SVM) machine learning method but this is for road vehicles. The research above have all done excellent work on solutions for specific problems on specific machines, as is the case for the CMAS solution discussed in this paper. However, no research was found utilizing machine learning for solving universal vibration analysis problems related to predictive maintenance for industrial machinery.

Except for machine learning, other state-of-the-art, high-quality research in vibration analysis is discussed in References 5, 12 and 21. The research[5] focuses on monitoring bearing condition in low- and medium-speed shafts by using the cepstrum, minimum entropy deconvolution, and kurtosis methods on preprocessing and feature extraction. The research[21] utilizes enveloping and principal component analysis (PCA) techniques to estimate the bearing condition and predict the bearing life. McDonald et al. in Reference 12 utilize a narrow band filter on different frequency bands to scan the time waveform signal and calculate the related kurtosis (for the purpose of deconvolution) to detect the impulse-like vibration which is associated with most bearing and gearbox potential faults.

Approach

Machine Learning Exception System for RDCA — The proposed machine learning exception system for RDCA has a schema per Fig. 1. The OSVM model was trained for each measurement location individually. Each measurement on the measurement location was defined as a measurement event (ME). The first step (pre-processing of raw data) excludes all data with a negative label on each ME so all remaining data is healthy data with a positive label. A measurement location with at least two years of historical data typically have more than 20 positive MEs. However, 20 data points are not enough for training a model. So, a search had to be conducted over a larger data set consisting of 214,303 signals from eight different large-scale and similar mills to find 100–200 spectra with similar morphology for each of the original 20 MEs. In this case, the search resulted in providing between 2,000 and 4,000 additional data points for training, which was sufficient to enable the OSVM to converge and guarantee ample training data set diversity.

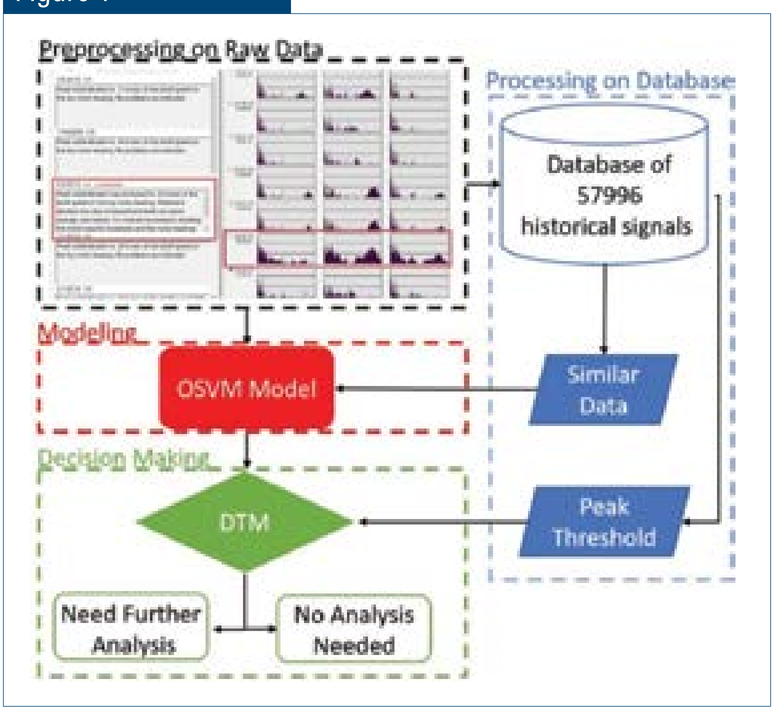
Search Method Based on Correlation: According to the unique characteristics of data analysis on vibration signals, both the amplitude and morphology of the spectrum are very important, e.g., unbalance, misalignment and looseness usually exhibit as maximum peak amplitudes rising in the spectrum. However, some early stage bearing issues will not show as one of the highest peaks, but will change the morphology of spectrum as a rising high-frequency noise floor. Consequently, ITR used a searching method to find the signals in the database with similar peak amplitude and morphology. An error tolerance T_e was set on the maximum peak amplitude and used to define the maximum acceptable tolerance, according to Eq. 1, where P_t is the maximum value and R_p is the maximum value range.

$$R_p = P_t \cdot (1 \pm T_e)$$

(Eq. 1)

Utilizing the correlation defined in Eq. 2, ITR calculated the morphology similarity r between t and signals d.

Figure 1



System schemata for route-based data collection and analysis (RDCA).

Figure 2

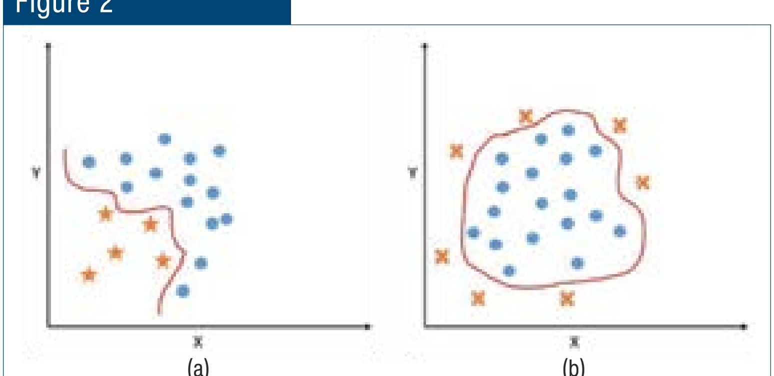


Illustration of support vector machine (SVM) (a) and one-class SVM (b).

$$r = \frac{n\left(\sum td\right) - \sum t \sum d}{\sqrt{\left[n\sum t^2 - \left(\sum t\right)^2\right]\left[n\sum d^2 - \left(\sum d\right)^2\right]}}$$

(Eq. 2)

A threshold T_r was set. If the signal had its maximum value in the range R_p and has an r larger than T_r, it met the requirement as candidate training data.

OSVM Machine Learning Classifier: As is generally known, a beneficial characteristic of binary SVM is that it can create a non-linear decision boundary by projecting the data through a non-linear function ϕ to a space with a higher dimension. The data points are lifted from the original space I into a feature space F at a higher level where a straight hyperplane can divide the data into classes. This straight hyper-plane can then be projected back to original space I to form a non-linear curve as Fig. 2a. The hyperplane separated by the two classes can be represented by $\omega^T + b = 0$, $\omega \in F$ and $b \in R$. The decision function (classification) rule for a data point x is:

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{n} a_i y_i K(x, x_i) + b\right)$$

(Eq. 3)

$$K(x, x_i) = \exp\left(-\frac{\left\| x - x_i \right\|^2}{2\sigma^2}\right)$$

(Eq. 4)

In machine learning, one-class classification tries to identify objects of a specific class amongst all objects primarily by learning from a training set containing only the objects of that class.[15] The most obvious difference between binary SVM and OSVM is the data for training. During the training of binary SVM, data with positive labels and negative labels are given (e.g., in Fig. 2a, the positive data are blue circles and the negative data are red stars). For training the OSVM, only one-class data with positive labels are given and an outlier detection function MAD in Eq. 5 needs to be processed on the data to sacrifice a small amount of data as outliers (the red crosses in Fig. 2b).

The outliers can be used as negative data and a SVM can be fit to the data as the regular binary SVM. A decision boundary is formed as the red circle in Fig. 2b. The boundary can be expanded or contracted accordingly by tuning the fraction between the outlier and desired data. The same decision functions shown in Eqs. 3 and 4 are used to calculate whether new data points are out of the boundary (the data may indicate machine fault) or inside the boundary (the data are OK and do not need a further analysis).

$$MAD = median(|A_i - median(A)|)$$

(Eq. 5)

Decision-Making Combined With Traditional Threshold: Unique decision-making schemes are applied separately to the velocity spectra and acceleration spectra. For the velocity spectra, analysts are more concerned with both the maximum peak value at multiples of the shaft speed as well as the harmonics occurring throughout the spectra. Harmonics directly affect the morphology of the data. The decision-making for velocity spectra checked the statistical peak values extracted from historical positive labeled data using Eq. 6, in which T_s is the threshold for peak values, μ is the mean value extracted from historical "OK" spectra, and σ is the related standard deviation. OSVM was also used to evaluate the morphology. For the acceleration spectra, analysts are more interested in high-frequency peaks (which may be low in amplitude) and the high-frequency noise floor. Potential faults exhibit more in the shape of the acceleration data than the highest peaks. This is the reason only OSVM is used.

If the algorithms using OSVM return a result showing morphology change or a peak value exceeding the statistical threshold in Eq. 6, the current velocity spectrum under testing is classified as "Need Further Analysis."

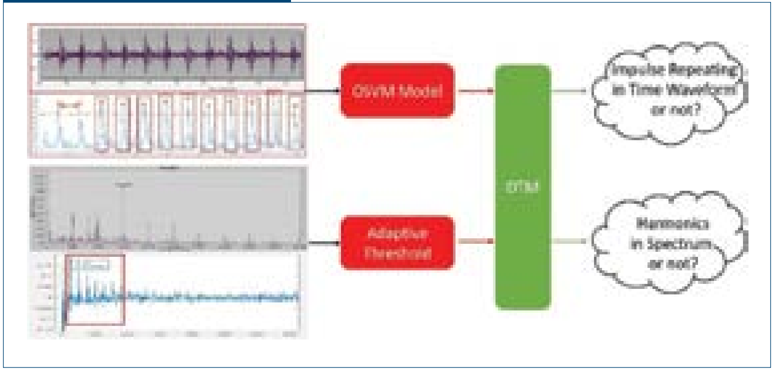
$$T_s = \mu + 2\sigma$$

(Eq. 6)

The algorithms evaluating the acceleration spectra only consider the OSVM result for morphology change because the critical peak values of concern are often low amplitude but have a raised high-frequency noise floor and harmonics across the full spectrum, both of which may indicate a potential issue. Both of these cases have a dramatic change on spectrum's morphology and this is detected by OSVM.

Using Semantic Analysis to Purify the Training Data That Can Be Expanded or Contracted: As a practicality of being a predictive maintenance service provider, ITR experts frequently label machines as "OK" but indicated some minor issues in their report; e.g., "The harmonics on inboard motor increased, but no immediately problem. Let's have a look at the next reading." This type of reporting is valuable to the customer because it notifies them of future potential problems, it is being trended, and no action is yet necessary. But for the purpose of training machine learning models, the associated data with this reporting usually has some signatures to machinery faults and can lead to "minor"

Figure 3



System schemata for condition monitoring and analysis systems (CMAS).

Figure 4

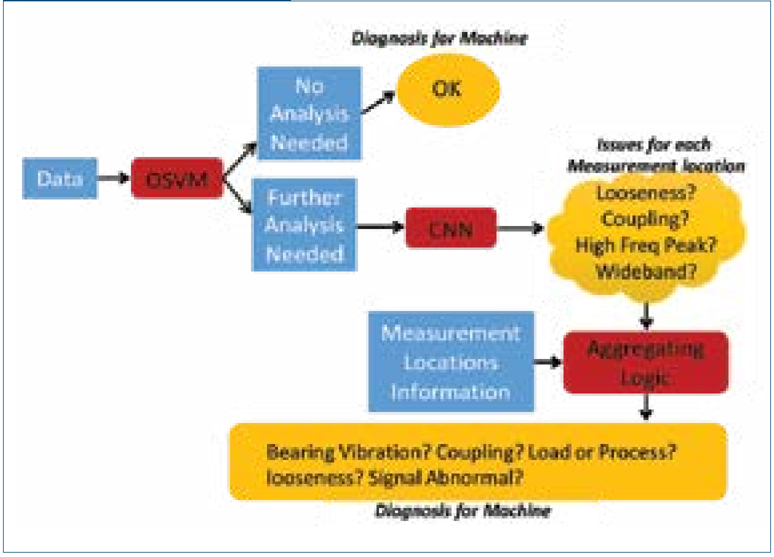
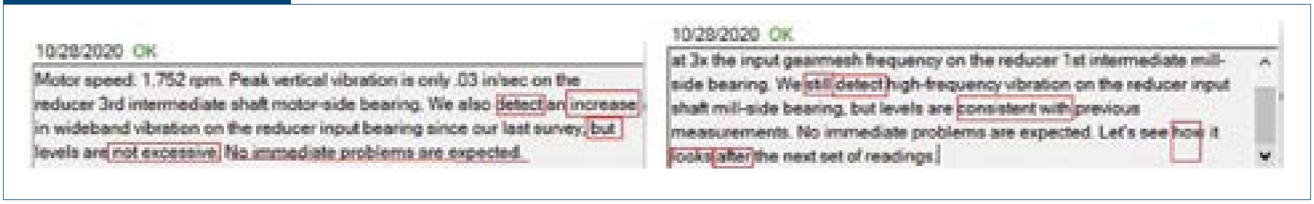


Illustration of convolutional neural network (CNN) further classification and aggregating logic.

Figure 5



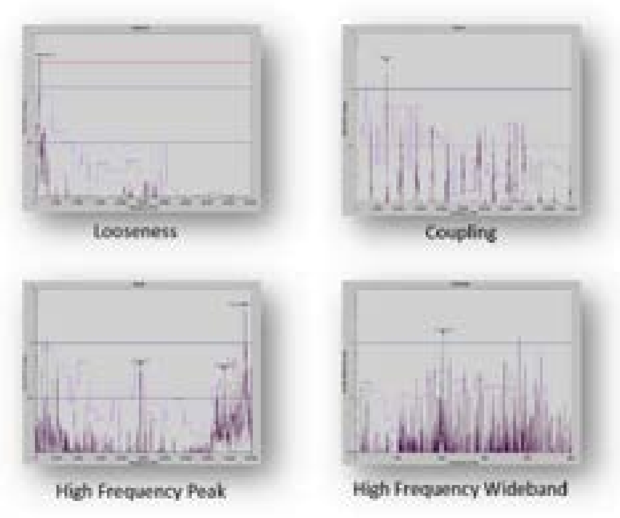
Using semantic analysis model identifying Not 100% OK Data.

false positives. In these cases, a machine learning technique was utilized to design a semantics analysis model. This model can process the experts' reports and classify them into "100% OK" data and "Not 100% OK" data. Fig. 5 illustrates how the semantics analysis model captures the sensitive phrases and delineates between 100% OK and Not 100% OK data from the language of prior analysis from experts. During the training or retraining for the OSVM models, these Not 100% OK data will be excluded out in the training set. This approach significantly increased the accuracy of the machine learning exception system and boosted the efficiency of the training or retraining process while still avoiding type II errors.

Convolutional Neural Network Further Classification: CNN, a class of deep neural networks, is very common in analyzing visual imagery. Based on the shared-weights architecture and translation invariance characteristics,[22] CNN has applications in image and video recognition, recommender systems,[20] image classification, medical image analysis and natural language processing.[4]

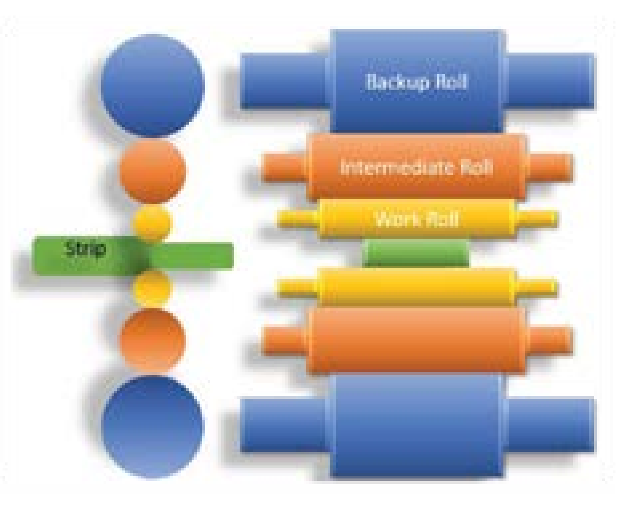
Based on CNN, the function of the machine learning exception system was expanded. In addition to classifying data as Need Further Analysis or No Analysis Needed, the system now also provides information on the nature of identified issues for each measurement location (ML, see Fig. 4). By leveraging the advantage of CNN for image classification or recognition, the issues within each spectrum can be always identified, no matter how relevant feature data moves (left or right), twists or is magnified. Where previously each ML used separate OSVM models, all MLs use one general CNN model for signature recognition and issue identification. Four types of issues are currently identifiable with this method: machine or component looseness (high-shaft-speed peak or harmonics), coupling issues (high axial vibration, including issues related to misalignment), high-frequency peak issues (electrical or load-related problem, some gear mesh-related issues, lobe pass issues, etc.), and high wide-band vibration (raised noise floor in the high-frequency range related to mid- to late-stage bearing issues) (see Fig. 6). With the data associated with each ML and asset (e.g., motor, fan, compressor,

Figure 6



Signatures view.

Figure 7



Pickle line tandem cold mill.

velocity or acceleration spectrum, etc.), a logic algorithm was developed to aggregate the issues identified for each measurement location and generate a diagnosis for each machine with severity (one star, two stars or three stars, which represent the minor issue, the moderate issue, and the severe issue, respectively). The machine's diagnosis also has one or more labels: bearing vibration, coupling, load or process issue, looseness, and abnormal signal. The aggregating logic is based on the results from prior expert analysis. When multiple different signatures are identified on various measurement locations, the algorithm evaluates each signature's severity and outputs up to two of the most severe issues. Using the aforementioned semantic analysis technique, the exception system also generates a simple report for each machine.

Confidential Score Generation: A customized Sigmoid function (Eq. 7 and Eq. 8, in which x is the classification scores from OSVM and n is the parameter to adjust the slope of the curve in Sigmoid) was developed to apply a confidential score to the exception system. The higher the score is, more likely the ML is properly categorized as Need Further Analysis.

$$S(x) = \frac{e^x}{e^x + 1}$$

(Eq. 7)

$$Score = 2 \cdot (1 - S(x/n))$$

(Eq. 8)

By implementing a threshold on these confidential scores, it allowed for the sensitivity of the machine learning exception system to be manually adjusted.

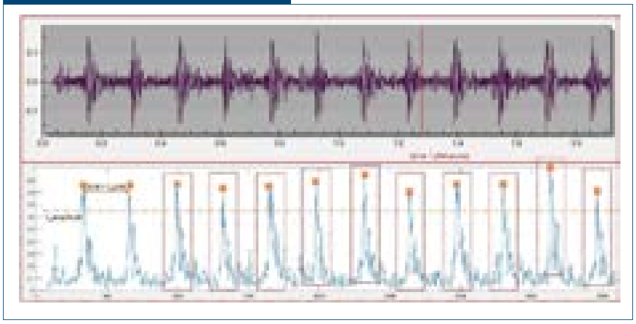
Impulse and Harmonics Detection for CMAS — A CMAS has both advantages and disadvantages over an RDCA. The advantage is considerably more data; in this test case, data is collected every few minutes continually. The disadvantage is the lack of verified expert analysis correlated to the majority of the data. Traditional feature thresholds such as RMS, peak, and crest factor are used and are effective when the fault is already in late stages or is well developed. But this approach often fails to detect minor faults in early stages or faults on variable load and speed machines. It can also deliver false positives. Accurate early detection is very important to maintenance and reliability professionals to enable preventive and corrective maintenance planning. According to References 3, 11 and 19, impulses in time waveforms and harmonics in spectra are good indicators of early minor machine faults. For the PLTCM CMAS, impulses repeating in the time waveform and harmonics in spectrum relate to spalls on the work and backup rolls and early bearing issues. These issues result in fifth-octave chatter, causing unacceptable product quality. As shown in Fig. 3, by leveraging the customized feature extraction for impulse and cepstrum (OSVM and DTM), the method deployed to the CMAS on the PLTCM stands detected the impulses in the time waveform and the harmonics in the spectra accurately and without any false alarms. This was verified over 12 months of testing in 2019–2020.

As the structure in Fig. 7 shows, the PLTCM is a mill system consisting of three to six mill stands, each stand consisting of work rolls, intermediate rolls and backup rolls arranged in series to progressively reduce the thickness of the strip in a single pass. Most products produced by a PLTCM are cosmetic and surface quality is extremely important. The PLTCM is extremely sensitive to changes in vibration since unacceptable vibration can cause surface defects resulting in product that will not be accepted by the customer.

Due to the complexity and the overall geometry of the PLTCM, the best sensor placement available for the system and objectives is on the top of the stand. The accelerometer senses the vibration from the three sets of rolls on the stand (work, intermediate and backup rolls) and the bearings for each roll. The PLTCM operates as a variable load and speed mill and produces several different product types. Vibration amplitudes and frequencies vary greatly. Consequently, the setting of thresholds for RMS, peak and crest factor are very difficult. On the other hand, the PLTCM structure is very sensitive and interconnected. Because of the compact positions of the work, for intermediate and backup rolls in series, if one of them develops a fault, a vibration may transmit to other parts and lead to more severe machine failures in a short time and adversely affect product quality. So, it is important to continuously monitor the impulses in the time waveform (related to chatter) and harmonics in spectrum (related to bearing issues, also potentially influencing chatter).

Customized Repeating Impulse Detection for Time Waveform: Utilizing Pan Tompkins peak detection method in Reference 16, the spikes in any time waveform can be easily located. However, since the algorithm is too sensitive, the spikes of the noise floor are also marked. Most of the unwanted spikes are filtered out with the enveloping technique and the customized dynamic thresholds set on the averaged value of n maximum values in time waveform and the minimum distance between adjacent spikes calculated from the rolls' speeds in Eq. 9 (in which v_w is the speed of work roll, v_i is the speed of intermediate roll, v_b is the speed of backup roll, f_s is the sample rate for the target time waveform).

Figure 8



Impulse detection on time waveform.

$$T_{dist} = \frac{f_s}{\max\left(u_w, u_i, u_b\right)}$$

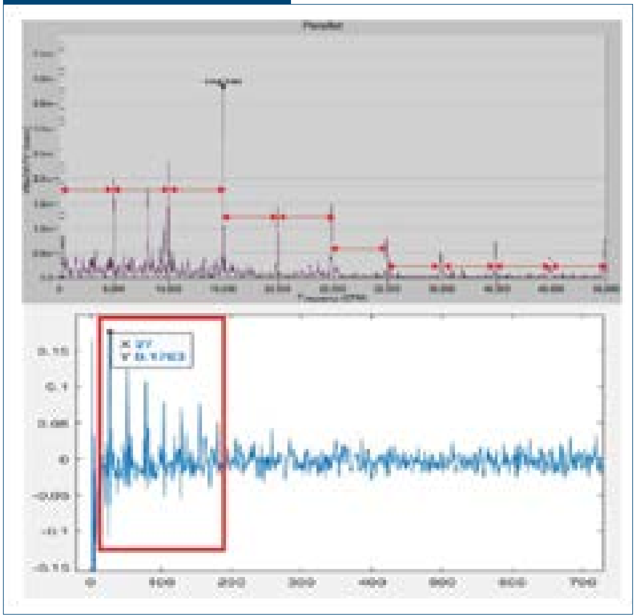
(Eq. 9)

The picture at the top of Fig. 8 is the original time waveform and the picture at bottom is the normalized signal after enveloping. With the dynamic threshold $T_{amplitude}$ set on spike amplitude and the threshold T_{dist} set on minimum distance between adjacent peaks, these repeating spikes are located accurately. Through a window of $loc_{spike} \pm 1/2$ in which loc_{spike} is the spike's location and l is the window length, the impulses can be defined individually. The crest factors ($CF = x_{peak}/x_{rms}$) for each window can also be calculated and averaged to get an averaged CF_{avg}. By setting a threshold on this averaged crest factor, most of the impulses that are not relevant are eliminated.

The standard deviation of the distances of every two adjacent spikes, the number of spikes and the averaged raw signal after normalization in each window were extracted as features for training the machine learning OSVM. Repeating impulses were detected if the OSVM returns a positive (the positive is different here from the one in the Machine Learning Exception System for RDCA section; because the algorithm used the time waveform with repeating impulses to train this OSVM, the positive result means the OSVM classified the data into "repeating impulses").

Cepstrum-Based Harmonics Detection Method: The harmonics of interest are visible in the top picture in Fig. 9. Although the overall amplitude is very small (because the measurement point is far away from the vibration source, the maximum peak has only 0.004 inch/second, or 0.102 mm/second), the harmonics are still seen very clearly. These harmonics usually relate to the roll bearing defects (work roll, intermediate and backup roll shown in Fig. 7). Traditional thresholds on RMS and crest factors did not work well enough for identifying this "harmonics" issue. However, cepstrum can combine and concentrate the energy of these harmonics, resulting in differentiated "harmonics" and "not harmonics" by threshold. Cepstrum is defined as follows: the squared magnitude of the inverse Fourier transform of the logarithm of the squared magnitude of the Fourier transform of a signal:[1,14] $Cepstrum = \left|F^{-1}\{\log(\left|F\{f(t)\}\right|^2)\}\right|^2$. First, the signal is normalized. After normalizing, the log of the spectrum is calculated, which makes the signal more periodic (more like a sinusoidal/cosine signal). And the inverse FFT projects the logarithmic signal back to a signal in the cepstrum's time series and the harmonics tend to concentrate into one or several high spikes, as the picture at the bottom of Fig. 9 shows. However, the non-linearity of the logarithmic signal sometimes will lead to more than one spike in the projected signal. To further increase the accuracy of this harmonics detection method, an adaptive threshold scheme was applied for relating these spikes' amplitudes to bearing defects. Using an equation similar to Eq. 9, the window size of the red rectangle in Fig. 9 was determined. In the window, two different thresholds for the maximum spike amplitude and the sum of the spike and its harmonics (the spikes at 2x, 3x, etc.) were set. If either the dominating spike or the sum of this spike and its harmonics exceeds one of the established thresholds and the spike's location is relevant to one of the bearing defect frequencies, the system identifies a potential issue. The calculation to determine if there is relevance is Eq. 10. β is the function for bearing defect frequencies given the dimension of the roll and machine speed (e.g., diameter of rolls, diameter of bearing rollers, number of bearing rollers, strip speed, motor speed, etc.). d_{work}, d_{int} and

Figure 9



Harmonics and corresponding cepstrum.

d_{back} are dimensions of work roll, intermediate roll and backup roll, respectively. υ is the motor speed. f_s is the sampling rate of time waveform. l is location of the spike at front in the cepstrum:

$$\frac{f_s}{l} \in \{b(d_{work}, \upsilon), b(d_{int}, \upsilon), b(d_{back}, \upsilon)\}$$

(Eq. 10)

Algorithm 1

Input: S represents the trust score, T represents the alarm threshold, S_{min} represents the minimum of the score, S_{max} represents the maximum of the score, R is the result from OSVM or adaptive threshold system

Output: *alarm_or_not* (*True* or *False*)

1 Set the initial trust score $S \leftarrow S_{max}$;
2 **if** $S > S_{min}$ and R is abnormality **then**
3 | $S--$;
4 **else if** $S < S_{max}$ and R is normality **then**
5 | $S++$;
6 **if** $S < T$ and R is abnormality **then**
7 | *alarm_or_not* ← *True*;
8 | Alarm is raised, system will send an notification to human analyst;
9 **else**
10 | *alarm_or_not* ← *False*;
11 **end**
12 Go back to Step 2 with the updated R on the new coming in signal;

Dynamic Trust Model: Because of the continuous monitoring characteristics of the CMAS and the importance of avoiding false alarms, both issues of repeating impulses or harmonics will lead to a repeating abnormality, meaning more than one abnormal signal will keep showing up uninterrupted, and this situation will always exist before the machine fault is corrected. In this case, for the purpose of minimizing false alarms, the algorithm used a fluctuating trust score that adjusts according to the signal detected as described in Algorithm 1; e.g., if the signal is abnormal and has repeating impulses or harmonics, the score will drop; if not, the score will climb or stay at its maximum level. An alarm/notification was sent out only if the score dropped below the threshold. This approach ensured no false alarms were given and the timeliness delays (from the DTM evaluation) were negligible. This trust score system (DTM) has already been broadly used in continuous authentication systems based on biometrics.[2,9,10,13]

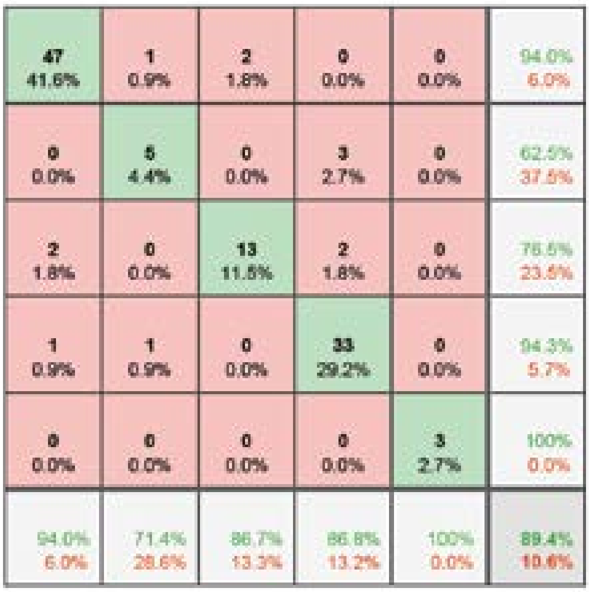
Evaluation and Result

Performance of Machine Learning Exception System for RDCA – For RDCA, 17,962 OSVM models were trained on 17,962 measurement location points across several

Figure 10

Performance for identifying "OK."

Figure 11



Performance for classifying diagnosis.

mills with more than two years of historical data. The exception system generated two possible results on testing spectra (0–833 Hz for velocity spectra and 0–4,167 Hz for acceleration spectra). After going through the aggregating logic, the possible diagnosis for each of 1,932 machines include:

- OK: The signal from all measurement locations on this machine has a high similarity on both morphology and amplitude with some of the historical positive labeled (OK) data, and the machine does not need further analysis by an expert.
- Need Further Analysis: The one or more measurement location signals from the current machine do not match with any of the historical positive labeled data, has a very high possibility of indicating a potential machine fault, and needs further analysis by an expert. If the machine is identified as Need Further Analysis, it will be further diagnosed as having one or more of the following conditions:
  - Bearing Vibration: Higher noise floor or obvious harmonics related to bearing defects present in acceleration measurement locations. Recommendations often include lubricating or replacing the bearing, depending upon severity.
  - Coupling: It is also called Misalignment and often lets the velocity measurement location signal have higher shaft speed peak and more obvious shaft speed peak's harmonics in the axial direction. It is usually suggested that the customer check the alignment and correct the coupling issue.
  - Load or Process: Characteristics include higher gear mesh, lobe pass or fan pass in non-motor measurement locations. When severity is minor, recommendations include decreasing machine loading, when possible. When severity is higher than moderate, recommendations include detailed inspections and possible repairs.
  - Looseness: Similar to coupling issues, however, characteristics include higher shaft speed peaks and harmonics present in radial directions but not in the axial direction. Recommendations include checking for soft foot, loose bolts or other looseness-related corrections.
  - Signal Abnormal: Related to quality of data, the system's pattern recognition capabilities include identifying sensor and cable issues that indicate bad data. Recommendations include remeasurements or repair/replacement (RDCA) or replacement (CMAS).

Fig. 10 presents the machine learning exception system's performance for identifying "OK" and "Not OK." During the period of December 2020 to February 2021, 956 machines were tested. Among the 729 OK machines, the exception system outputs only misclassified one Not OK machine as OK. In this case, the true positive rate for predicting OK machines is 99.9%. The one misclassified machine here had a minor Looseness issue, which is an event on the margin between OK and Not OK. Additionally, the exception system output 227 machines as Not OK. Only 113 were truly Not OK, which results in a true negative rate of 49.9%. However, considering the situation that data analysts often indicate some minor issue in the report but label the machine as OK (mentioned in the Using Semantic Analysis to Purify the Training Data section) and the objective is designing an exception system with a 100% true positive rate, the relatively low true negative rate is acceptable. Summarizing, among the 956 machines tested, the machine learning exception system labeled 728 (76.15% of all) machines as OK with a 99.9% accuracy.

Among the 113 Not OK machines, the exception system diagnosed them further and assigned condition codes. Fig. 11 shows the system performance results. The exception system outputted 50 Bearing Vibration classifications; 47 of them are classified correctly. One machine was misclassified into the Coupling category, and two machines were misclassified into Load or Process. There were eight Coupling classifications from the exception system; five of them are classified correctly and three of them were misclassified as Looseness. Among 17 Load or Process classifications, four were misclassified. There were 35 Looseness classifications; all but two were correctly classified. Finally, there were three Signal Abnormal classifications, all without any error. Combining the results, the rate of correct classification was 89.4%.

This machine learning exception system is implemented in production environments and performing well. Fig. 12 shows output from a future artificial intelligence (AI) exception system currently in development and testing that takes the production system one step further. This system uses the severity determination (none, one star (minor issue), two stars (moderate) or three stars (severe)) and diagnosis to also generate an expert report based on historical findings language. As a representative example of the current results from the system, the left column in the figure is the output of the AI exception system, and the right column is the report and diagnosis of ITR experts. Testing has shown the system generates similar diagnoses and correctly identifies the asset components and associated measure locations but with fewer details. As designed, it tends to diagnose more issues with higher severity. Based on early performance, further development is ongoing and it is

Figure 12

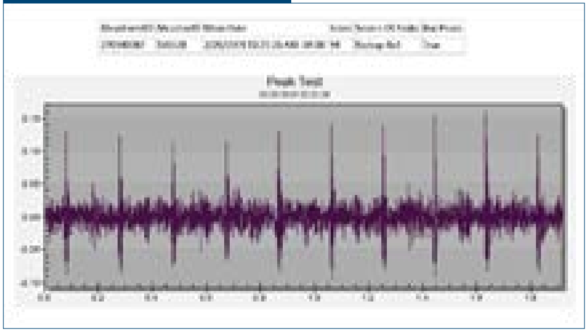
Report Generated By Machine Learning Exception	Report Written By Data Analyst
(2-star, Looseness) Peak Vertical vibration is 0.65 ips on Outboard Motor Bearing (407481). This is high. Peak Axial vibration is 0.2 ips on Inboard Motor Bearing (407482). It is slightly above average.	(1-star, Looseness) Peak vertical vibration has slightly decreased from .70 in/sec to .65 in/sec at the shaft speed on the outboard motor bearing. This is still moderate vibration, but levels are near the long-term average. No immediate problems are expected.
(1-star, Bearing Vibration) High Frequency vibration is above average on Outboard Motor Bearing (610211).	(OK) We still detect high-frequency shaft speed harmonics on the motor bearings, but vibration levels are stable. No immediate problems are expected.
(OK) Peak Horizontal vibration is 0.04 ips on Motor. No problems are indicated.	(OK) Peak horizontal vibration is .04 in/sec on the motor. No problems are indicated.

Report generation function exhibition.

expected this additional system will be deployed to production environments in the near future.

Performance of Impulse and Harmonics Detection for CMAS — Due to the tremendous amount of data collected by a CMAS, it is impossible for experts to analyze every spectrum, waveform and feature set collected and processed as was the case with the RDCA. However, the CMAS has its own notification system. If any monitored feature parameters, e.g., overall RMS, banded RMS, peak values or crest factor, exceed the pre-set threshold, it sends a notification to the expert (or any interested party). As mentioned in the Impulse and Harmonics Detection for CMAS section, the threshold on these signatures do not work well on the early-stage problems related to repeating impulse and harmonics. Leveraging the unique feature extraction, the machine learning classifier (the OSVM and DTM system) and the impulse and harmonics detection program, the CMAS successfully detected 19 bearing issue and 41 spall issues related to chatter over 24 months, as shown in Table 1. The bearing issue was diagnosed from harmonics within the spectrum, and the spalling issues were diagnosed from repeating impulses

Figure 13



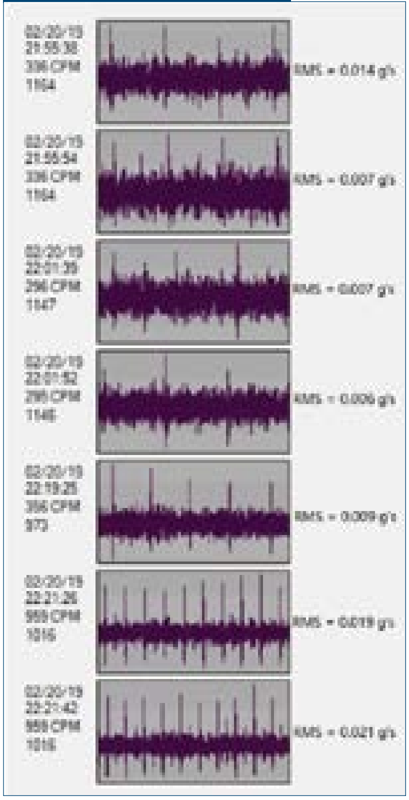
Impulse detection notification example.

Table 1

Impulse and Harmonics Detection Result on PLTCM for January 2019–March 2021

No. of measurement events tested	Harmonics detected (bearing issue)	Impulse detected (spalls)	False alarms	Missed problems
315,988	19	41	2	0

Figure 14



Repeating impulses example.

in the time waveform. From January 2019 through March 2021, 315,988 MEs were tested. The model operated successfully without any missed problems and detected many harmonics events and impulse events for all five stands of the tandem mill within the PLTCM. Two false alarms were generated over the 24 months, but after investigation, these false alarms were attributable to problems with damaged sensor and cable connections, resulting distorted signals.

Fig. 13 is an example showing one of the notifications sent to the expert for repeating impulses. The notification was sent out on 20 February 2019 at 10:21:26 p.m. EST on measurement location 165528 #2 stand mill drive on PLTCM with a DTM score of 94, which is lower than the alarm threshold and the source of impulse is the backup roll. The plot for the time waveform in the notification email has very obvious prominent repeating impulses. The expert who received the email notification analyzed the data and verified that the finding was correct. Fig. 14 shows the repeating impulses beginning in the time waveform at 9:55:38 p.m., but the machine was running at a relatively low speed of 295–336 rpm and the overall RMS is 0.007–0.014 Gs. This initial impacting had already dropped the score in the DTM but it was still above the alarm threshold. At 10:21:26 p.m., the machine was running at a high speed of 959 rpm and the impulses were causing both RMS and the crest factor to increase. At this time, the uninterrupted time waveform with repeating impulses caused the DTM score to drop below the alarm threshold. This initiated an exception notification. Once the expert confirmed the finding, the mill was immediately notified of the issue, including the specific source of the problem on the backup roll. Instead of continuing to run 40–50 additional coils with chatter marks before inspection, the mill was able to stop, change the backup rolls and only scrap a partial coil.

The success of the spall detection relies on both impulse and harmonics detection by the CMAS and the experienced expert's quick reaction. This automatic machine learning detection system integrated with an expert analyst makes ITR's system truly capable of very minor false alarms, 100% detection rate and quick notification.

Future Work

The work described in this paper has achieved the intended objectives of creating a multi-layer scheme for classifying general machine condition and further classifying conditions, severity, and recommended actions. Based on the positive results, ITR will further test, refine, expand and deploy these models. Additionally, the new models will include data mining, deep learning techniques, and unsupervised machine learning feature extraction (e.g., Autoencoder, PCA and independent component analysis) to find the hidden features for these subsets. This will not only improve response times, but it will also help improve the precision of recommended actions.

For the impulse and harmonics detection for the CMAS, ITR has already realized a customized method on the PLTCM and it continues to return excellent results. In the future, ITR will expand this method to other installed CMAS and incorporate this technology into new system installs. ITR has begun work to extend these models to other asset types, tailored to the potential failure modes specific to the particular operation.

Lastly, the machine learning exception models developed for RDCA will be refined for CMAS. DTM will be used to boost the true negative rate. The CMAS exception system will capture abnormal measurement events and send alarms (exception notifications) with diagnosis and severity suggested by AI to experts or operators in real-time to realize a 24/7 and more precise monitoring and diagnosis.

References

1. B.P. Bogert, "The Frequency Analysis of Time Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum and Phase Cracking," *Time Series Analysis,* 1963, pp. 209–243.
2. P. Bours, "Continuous Keystroke Dynamics: A Different Perspective Towards Biometric Evaluation," *Information Security Technical Report 17,* No. 1, 2012, pp. 36–43.
3. H. Chen, P.S. Chua and G. Lim, "Fault Degradation Assessment of Water Hydraulic Motor by Impulse Vibration Signal With Wavelet Packet Analysis and Kolmogorov-Smirnov Test," *Mechanical Systems and Signal Processing,* Vol. 22, No. 7, 2008 pp. 1670–1684.
4. R. Collobert and J. Weston, "A Unified Architecture for Natural Language Processing: Deep Neural Networks With Multitask Learning," *Proceedings of the 25th International Conference on Machine Learning,* 2008, pp. 160–167.
5. C. Freitas, P. Morais, J. Cuenca, A.P. Ompusunggu, M. Sarrazin and K. Janssens, "Condition Monitoring of Bearings Under Medium and Low Rotational Speed," *European Workshop in Structural Health Monitoring,* 2016.
6. M. Gan and C. Wang, "Construction of Hierarchical Diagnosis Network Based on Deep Learning and Its Application in the Fault Pattern Recognition of Rolling Element Bearings," *Mechanical Systems and Signal Processing,* Vol. 72, 2016, pp. 92–104.
7. Y. Gao, W. Wang, B. Li, O. Patil and Z. Jin, "Replicating Your Heart: Exploring Presentation Attacks on ECG Biometrics," *2018 IEEE Proceedings on Communications and Network Security,* IEEE, 2018.
8. J. Lepine, V. Rouillard and M. Sek, "On the Use of Machine Learning to Detect Shocks in Road Vehicle Vibration Signals," *Packaging Technology and Science,* Vol. 30, No. 8, 2017, pp. 387–398.
9. B. Li, H. Sun, Y. Gao, V.V. Phoha and Z. Jin, "Enhanced Free-Text Keystroke Continuous Authentication Based on Dynamics of Wrist Motion," *2017 IEEE Workshop on Information Forensics and Security (WIFS),* IEEE, 2017, pp. 1–6.
10. B. Li, W. Wang, Y. Gao, V.V. Phoha and Z. Jin, "Hand in Motion: Enhanced Authentication Through Wrist and Mouse Movement," *9th IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS),* IEEE, 2018, pp. 1–8.

11. B. Luo, H. Wang, H. Liu, B. Li and F. Peng, "Early Fault Detection of Machine Tools Based on Deep Learning and Dynamic Identification," *IEEE Transactions on Industrial Electronics,* Vol. 66, No. 1, 2019, pp. 509–518.
12. G.L. McDonald, Q. Zhao and M.J. Zuo, "Maximum Correlated Kurtosis Deconvolution and Application on Gear Tooth Chip Fault Detection," *Mechanical Systems and Signal Processing,* Vol. 33, 2012, pp. 237–255.
13. S. Mondal and P. Bours, "A Computational Approach to the Continuous Authentication Biometric System," *Information Sciences,* Vol. 304, 2015, pp. 28–53.
14. M.P. Norton and D.G. Karczub, *Fundamentals of Noise and Vibration Analysis for Engineers,* Cambridge University Press, 2003.
15. P. Oliveri, "Class-Modelling in Food Analytical Chemistry: Development, Sampling, Optimisation and Validation Issues — A Tutorial," *Analytica Chimica Acta,* Vol. 982, 2017, pp. 9–19.
16. J. Pan and W.J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomed. Eng,* Vol. 32, No. 3, 1985, pp. 230–236.
17. B. Schölkopf, A.J. Smola, F. Bach, et al., *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond,* MIT Press, 2002.
18. J. Shawe-Taylor and N. Cristianini, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods,* Vol. 204, Cambridge University Press Cambridge, 2000.
19. J. Shiroishi, Y. Li, S. Liang, T. Kurfess and S. Danyluk, "Bearing Condition Diagnostics Via Vibration and Acoustic Emission Measurements," *Mechanical Systems and Signal Processing,* Vol. 11, No. 5, 1997, pp. 693–705.
20. A. Van den Oord, S. Dieleman and B. Schrauwen, "Deep Content-Based Music Recommendation," *Advances in Neural Information Processing Systems,* 2013, pp. 2643–2651.
21. T. Wang, "Bearing Life Prediction Based on Vibration Signals: A Case Study and Lessons Learned," *2012 IEEE Conference on Prognostics and Health Management,* IEEE, 2012, pp. 1–7.
22. W. Zhang, K. Itoh, J. Tanida and Y. Ichioka, "Parallel Distributed Processing Model With Local Space-Invariant Interconnections and Its Optical Architecture," *Applied Optics,* Vol. 29, No. 32, 1990, pp. 4790–4797. ✦

AISTech

This paper was presented at AISTech 2021 — The Iron & Steel Technology Conference and Exposition, Nashville, Tenn., USA, and published in the AISTech 2021 Conference Proceedings.