OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Original Article

A Qualified review of ML and DL Algorithms for Bearing Fault Diagnosis

Asma Bibi[1], Bushra Naz[2], Shahnawaz Talpur[2], Shahzad Hyder Soomro[2], Yusrah Bablani[3]
[1]Department of Computer Systems Engineering, Mehran
[2]University of Engineering and Technology, Jamshoro, Pakistan
[3]Department of Mechanical Engineering Case western Reserve University, Cleveland,Ohio
***Correspondence**: Asma Bibi, asma1997shah@gmail.com,

Citation Bibi.A, Naz.B, Talpur.S, Soomro.S.H, Bablani.Y "A Qualified review of ML and DL algorithms for Bearing Fault Diagnosis". International Journal of Innovations in Science & Technology, Vol 04 Issue 04: pp 998-1010, 2022.
DOI | https://doi.org/ 10.33411/IJIST/2022040401
Received |Sep 04, 2022; **Revised** | Oct 15, 2022; **Accepted** | Oct 22, 2022; **Published** | Oct 26, 2022.

Moving machinery is the backbone of socio-economic development. The use of machines help in increasing the production of everyday used items, and tools, that generate electricity and mechanical energy, and provides easy and fast transportation and help by saving human efforts, energy, and time. The mechanical industry is totally dependent on the bearing and it is considered as bread and butter of the system. Bearing failure is about 40% of the total failures of induction motors which is why it is a crucial challenge to predict the failure and helps prevent future downtime events through maintenance schedules with the latest techniques and tools of. This paper presents a review of how DL techniques and algorithms outsmarted ML for bearing fault detection and diagnosis and summarizes the accuracy results generated by most common DL algorithms over classical ML algorithms.Additionally this paper reasons different criteria for which DL algorithms have been proved efficient for building productive model in the field of bearing fault detection. Furthermore, some of the most famous datasets by different universities have been discussed and accuracy results are provided by reviewing algorithms on the CWRU dataset by different researchers and comparison chart is listed in the results section.

Keywords: Ball Bearing; CWRU dataset; Feature Engineering, Repetitive Neural Network; Generative model as Adversarial Network; convolutional neural network

Acknowledgment
The author is thankful to Mehran University of Engineering and Technology and its faculty from Computer Systems Engineering dept for providing guidance throughout the procedures. Special thanks to Madam Bushra Naz for wonderfully contributing as Supervisor. Plus, author is.Thankful to Case western Reserve University for publicly providing a dataset that made research easy and interesting

Author's Contribution
All authors have contributed equally.
Conflict of interest
The authors declare no conflict of interest in publishing this manuscript in IJIST.
Project details. NIL

Introduction

The word bearing itself is derived from the verb "to bear" which implies that bearing helps moving elements (balls or rollers) to bear and roll smoothly so that machines work tirelessly to keep downtime at the bay. Rotary machines based on the concept of "Things roll better than they slide" which are the backbone of the country's economy and industrialization should be functional [1]. At certain times these machines operate underneath negative conditions, which include high ambient temperature, moisture, and overload or electrical conditions like (a) Unbalance magnetic pull; (b) uncontrolled heat (c) Increased load (d) Efficiency reduction (e) Decrease in average torque (f) Enhanced torque pulsation, that may subsequently result in motor malfunctions thus leading to failure[2]. Failure of machinery at an unexpected time can cause heavy losses to the country's economy financially and technically due to the unavailability of healthy machinery to major sectors i.e., train system, wind turbine system, and industrial production machinery that result in high protection expenses, extreme economic losses, and protection [3]. Statistics show that more than 90% of machines use rolling components in industrial applications [4]. Bearing failure is about 40% of the total failures of induction motors that is why there is a crucial need to have a clear understanding of different algorithms in respect of datasets for a smart technology based autonomous online system that predict the failure and helps preventing future downtime events through maintenance schedules[5]. These autonomous systems are designed to help avoiding delayed customer orders on a largescale, waste of human resources, and overtime [6]. Different algorithms including Machine Learning(ML), Deep Learning(DL) under Artificial Intelligence(AI) have been used to predict fault and remaining life of a bearing [7]. Hence, these articles review the previous articles published on the rolling bearing in ML and DL with cooperation of AI in terms of technicalities and capacities in accordance with fault analysis to find out more accurate and efficient technique in predicting downtime for intelligent machine health monitoring [8]. The Figure 1(Parts of Ball Bearing) represents the outer race, an inner race which imprisons the central section of this simple machine i.e., bearing, a cage that maintains the distance between to bearing balls placed in connection that the other one next to the adjacent [9].

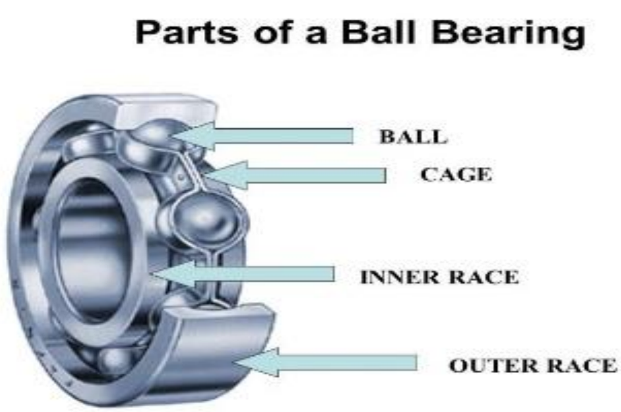

Figure 1: Parts of Ball Bearing

Bearing fault is the critical issue in a motor force the device, accurate diagnostics has become a matter of importance. Vibrations which are caused due to motion of a motor at certain speed, location of bearing at certain point while machines rotate. Since all these factors bring fluctuations in the amplitude of the vibratory signal that helps in predicting the fault [10][11]. A bearing with no-fault does not show any special change in generated vibration and as any fault is introduced in the outer race way normal sinusoidal wave is disrupted and at the peak point of a sine wave spike is observed which results in spectra and frequency increased and if it surpasses in the range of 20 to 20000k cycles per seconds it can be easily heard in a form of noise which is an early sign for fault detection in terms of manual detection but it does not help to measure the remaining healthy life of bearing or upcoming event which can be easily predicted by intelligent systemswhich are installed. Following figure 2 shows abnormality on the inner racetrack way[12]

Figure 2: High frequency detected in Malfunctioned Ball Bearing

In previous years Machine learning algorithms such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Artificial Neural Networks (ANN), and Convolutional Neural Networks (CNN) in the discipline of bearing fault diagnosis later interest grew in the field of Deep learning over time because of the reason that data collection grew more, and scientist predicted more accurate results in the field of bearing prognosis[13]. Fault detection can be treated as a classification problem and pattern recognition problem too. DL algorithms are potential tools for classification problems and researchers prefer this domain for condition checking and perceiving errors [14]. The fault detection may include sampling i.e., data acquisition as explained in Figure 3 which incorporates dealing with signal, data processing which extracts the time domain features of a signal such as a root mean square, skewness, kurtosis, and frequency domain features, feature extraction

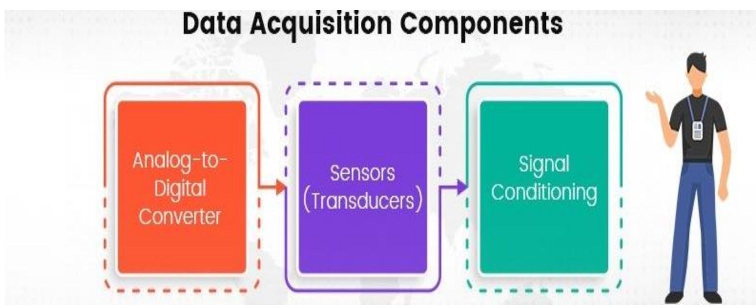

Figure 3: Digital Data Acquisition System

and then fault diagnosis incorporates the implementation of different algorithms [15]. Later Feature extraction algorithms help in preprocessing data so that input features can be generated and then algorithms are applied for the classification of faults. Usually, independent component analysis is used to convert the -dimensionalional vector low-dimensionalmensional vector for easier analysis [16]. The feature vector is used as an input vector for developing -enabledled fault identification system. algorithmsthms have become renowned in fault detection systems due to their reliability, adaptability, and robustness. The application of these algorithms has helped in developing an efficient system. These systems do not need prior knowledge for their operation [17]. The algorithms like (SVM), (k-NN), (ANN), random forest, (SVM), and (CNN) are commonly used in fault prognosis and diagnostic of bearing fault, and outputs are gathered and then reviewed. DL methods automatically learn rich features even from raw data using multiple computing processors together to achieve a better result [18]. Various DL architectures have been exploited for efficient bearing fault detection including Long Short-Term Memory (LSTM), Auto-Encoders (AE), Deep Belief Networks (DBN), (CNN) Long Short-Term Memory (LSTM). Among these DL models, CNNs vehave been the most representative modelhaved have demonstrated robust and remarkable performance as a supervised[19] [20].

A generic view

We require data for ML procedures and practices. It is considered a a crucial step thavinggood-qualitygood-quality dataset. Since natural degradation is sartificialy artificial faults are produced, and datasets are recorded for further experiments [21]. Some of the universities have published their datasets so engineers can use them and develop ML and DL algorithms. While in DL we don't need previously recorded information thus proving that DL algorithms have outsmarted most ML algorithms [22].

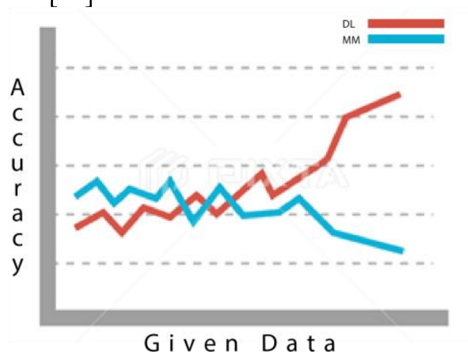

Figure 4: Performance comparison of ML over DL algorithms

In Figure 4 -axisis shows the Accuracy of DL Algorithms over ML algorithms over modern times whereas x declares how Deep learning algorithms show better accuracy over machine learning when the amount of data is increased [23][24] (Big dataset)

Objectives

1. Review existing research and datasets for bearing fault analysis and detection.
2. Comparing previous work done and finding out the most effective algorithm and method on CWRU **Dataset.**
3. Producing easy content guide for new researchers in the field of Bearing fault diagnosis and Prognosis.

Novelty:

Though GAN algorithms are optimized to provide higher fault accuracy than CNN (along with its modified type) and RNN (and its modified types) on datasets and are able to produce artificial faults mimicking the provided faulty set that means engineers do not need to utilize more resources to produce fault to study yet GAN is not a better option to apply due to a reason that it would not work on smaller datasets plus it is computationally expensive and it may not produce proper discrete data [25][26]. RNN is a better option available yet because it is capable of entertaining provided a heavy load and is able to memorize sequential events, but we frequently need to train [27]. RNN model which is the biggest drawback. CNN can extract a particular signal from the set of noise and signals combined but it needs to have a large, labeled dataset to work on [28][29].

Statement material and methods

APPROACHES FOR FEATURE LEARNING:

1. **Featurization For ML**

Featurization is a method to generate rows (a record) and columns (a feature that describe the record) from a given raw data i.e., text data, graph, or time series data. Featurization is the most pivotal step in ML done to gather relevant information as a feature vector as shown in Figure 5 or input to generate better results for predictive models or outputs manufactured by domain knowledge experts in contrast to supplying only the raw data to the machine learning process.

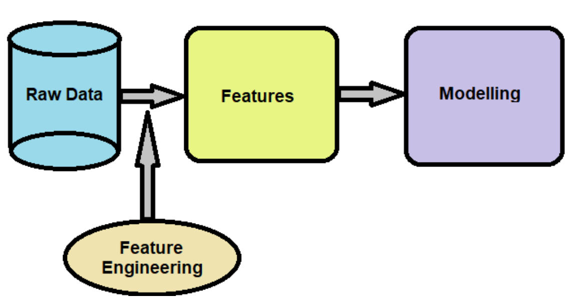

Figure 5: Feature Engineering for generating feature vectors/features

The feature engineering process includes the following techniques with further explanation in Figure 6.

1) Data preparation

In data preparation domain knowledge expert must refine n number of given datasets to integrate relevant data which may also refer to data cleaning of irrelevant rows and columns. The purpose is to feed meaningful and related data for true prediction.

2) Exploratory data analysis

When there is insufficient data provided exploratory data analysis helps reduce dimensionality between given features which brings out insightful features which would help reduce error rate and increase the predictability of the model under consideration.

3) Avoiding bias analysis

Bias is a systematic error that happens in machine learning itself due to incorrect assumptions in the ML process. Bias can be eliminated by using different analysis techniques over the same dataset in order to keep potential underfitting at the bay.

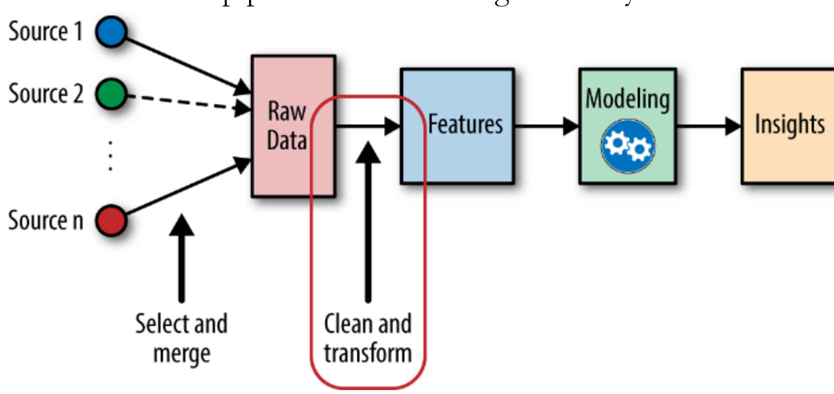

Figure 6: Feature Engineering Block Diagram [15]

2. Featurization with Deep NN

The DL algorithms are independent of human expertise along with the prior knowledge of the problem and have been observed to perform better in error identification, at times it is challenging to identify the faulty features precisely in presence of certain features such as noise. Methods specifically like DL from uncooked data and classing in a synchronized and tangled manner implement feature learning. When comparing, a lot of DL-related papers present outcomes by means of the classical-ML way with features, which are human-engineered for bearing error recognition

Most of the DL-related techniques are reported to leave behind the traditional methods, usually when uneven noise and recurrent variation in operation are present.

1. Thermal images.

Infrared thermography (IRT), warm video, and additionally warm imaging is a cycle where a thermal camera catches and producess a picture of an item by utilizing infrared radiation transmitted from the article in an interaction, which are instances of infrared imaging science. Thermographic cameras as a rule identify radiation in the long-infrared range of the electromagneticspectrum (around 9,000-14,000 nanometers or 9-14 μm) and produce pictures called thermograms. Since infrared radiation is produced by all articles with a temperature above outright zero as indicated by the dark body radiation regulation, thermography makes it conceivable to see one's current circumstance regardless of noticeable light. Thermography permits one to see varieties in temperature. At the point when seen through a warm imaging camera, warm articles stand apart well against cooler foundations; people and other warm-blooded creatures become effectively apparent against the climate, day, or night. Subsequently, thermography is especially valuable for monitoring a particullar Condition monitoring and applications Condition observing (informally, CM) is the most common way of checking a boundary of condition in hardware (vibration, temperature,, and so on), to recognize a huge changethatthat is demonstrative of a creating issue. It is a significant part of prescient support.

The utilization of condition observing permits upkeep to be planned, or different moves to be made to forestall weighty harms and keep away from its ramifications. Condition observing has a one-of-a-big advantage in that conditions that would abbreviate typical life expectancy can be tended to before they form into a significant disappointment. Condition checking procedures are typically utilized on pivoting hardware, assistant frameworks, and other apparatus (blowers, siphons, electric engines, -poweredered motors, presses), while intermittent assessment utilizing Non-Disastrous Testing (NDT) methods and Fit For Administration (FFS) assessment are utilized for static plant gear like steam boilers, channeling, and heat exchangers.

The following list includes the main condition-monitoring techniques applied in the industrial and transportation sectors:

i) Vibration analysis and diagnostics,
ii) Lubricant analysis) Acoustic emission
iii) Infrared thermography
iv) "Ultrasound"
v) Oil condition sensors
vi) (MCSA-Motor condition monitoring and motor current signature analysis)
vii) (MBVI- Model-based voltage and current systems).

Well known datasets

Case western reserve university (CWRU):

It is a very basic data set used by many the engineerings for ML and DL for finding optimal output. It uses 2 accelerometers as sensors and the sampling frequencies is 12 and 48 fault are generated artificially with a diameter of 7 mils, 14 mils,21mil, 28 mi, and 40 mils at the inner raceway.

Paderborn university dataset (PUD):

It uses total of 4 sensors which include 1 accelerometer 1 thermocouple, and 2 current sensors and the recorded frequency are 64 whereas the mode of fault production is artificial and accelerated aging. Basically, synchronous measurement of motor current and vibrational signals are used which help in merging distinct characteristics of signals which increases the probability of finding bearing fault.

Prognostic dataset (PD):

It is used for finding the remaining healthy and useful life of the rolling bearing element. The prognostic dataset includes 2 accelerometers, and 1 thermocouple, and the recorded sampling frequency are 25.6, and the mode of fault production in this dataset is natural. It provides real data. Intelligent maintenance systems database:

It requires 2 sensors as the accelerometers sampling frequency is 20 and the mode of fault induction in this data set is natural.

Dataset Intelligent Maintenance System (IMS):

IMS differs from the other datasets in such a way that faults in the bearing was created naturally by keeping the rolling bearing on the run for continuous 30 days with a constant speed of 2000 rpm.

It uses 2 accelerometers and the sampling frequency is 20 kHz and the mode of producing fault in the inner/outer race is natural.

Limitations of ML algorithms and reasons behind the usage of DL algorithms

To detect the fault in the bearing using conventional ML algorithms the characteristic fault frequencies are calculated (to train models to apply ML) which are generated by rotor mechanical speed and the placement of the bearing in any specific location. These generated frequencies would be used as fault features and determining these features would be termed feature engineering. ML algorithms can be applied at signal and changes are observed in amplitude if any irregularity is found.

Due to some reasons, deliberately introduced fault to bear faulty characteristic frequency that will be hard to understand leading to incorrect classification outcomes, specifically, when machine-learning procedures that depend on features, which are human-engineered during the preparation procedure. Hence, several DL systems that are self-computerized and extract features and use those features for better classification performance and bear faulty diagnostics, will converse in this review paper.

Case 1)

In some cases, the scientist accepts the shortcoming recurrence considering the thought that there was no sliding among the moving component and bearing raceway at the same time, frequently rolling and sliding are noticed. The created outcome is determined and can go amiss from the genuine issue recurrence.

Case 2)

In some other cases if multiple faults are assumed and the interaction is not justified in the sense that whether the faults have interacted or not and if any interaction is made what components have been generated and additional components are not added result would be unclear or somehow varying from the original result.

Case 3)

There are also the chances of external interference due to the placement of the element which can again interfere with the result.

Case 4)

Some features are sensitive to processed under conventional ML algorithms.

Reasons

The reasons for which the shift from ML to DL was observed

The main factors due to which the shift from ML to DL was observed, were based on performing practicals which concluded that classical ML algorithms can be performed better on a smaller datasets where are on larger data sets and the results are not efficient when such datasets were trained under DL the result can beat classical ML algorithms.

ML methodologies include manual feature engineering which is only possible with a person with great skills and expertise in the field and able to put the huge human effort in this domain but in deep learning the data is simply passed on a network and the system automatically learns the structures from raw records by auto-tuning the weightiness in the network.

Step comparison chart for Machine and Deep Learning Algorithms working	
Input	1. Input
Feature Extraction	2. Feature Extraction and classification
classification	3. output
output	-----------

Table 1 comparison chart of ML and DL procedures

Deep learning removes the challenging stage of feature engineering as shown in the right column of the table. DL-trained modules/systems are adaptive to any change provided such as the same convolution neural network can be used for visions and language problem-solving techniques.

Discussion

CNN: CNN (see more in CNN working module Figure:) is the most used algorithm nowadays and it was introduced by a scientist who got inspired by the concept that how humans understand things. Specifically, decrease layers within the network discover essential lower-degree visible features; and layers later stumble on better degree structures that are constructed on those easy lesser-degree structures. In 2016, the first paper employing CNN to perceive bearing fault turned into, and inside the subsequent 3years, numerous papers implementing the equal same have arisen and backed to evolving bearing faulty discovery in numerous factors. Especially, 1D temporal raw records acquired from unlike accelerometers are first stacked to 2D-vector method just like the representation of pics that is exceeded then above to a convolutional layer aimed at extracting the features, observed by consuming a pooling stage for downsampling. The combination of convolution-pooling samples is repetitive generally in addition to extend the community. The output from the hidden layers can be exceeded over to one or numerous completely linked layers, and the result of that is transferred to a pinnacle classifier based totally on SoftMax or Sigmoid capabilities to decide if a bearing fault is a gift...The vibration information is accumulated using uni-axis accelerometers hooked up on the x- and y- path respectively. A CNN can autonomously study beneficial functions for bearing fault detection from the uncooked facts pre-processed via scaled discrete Fourier transform. The category result demonstrates, the function mastering primarily based technique significantly outperforms the feature engineering-based method of traditional ML. Moreover, some other contributions of this painting are to expose that feature getting to know primarily based processes including CNN also can carry out bearing fitness prognostics and become aware of a few early-level faulty situations that have no specific function frequencies, inclusive of lubrication degradation, can't be accomplished the usage of classical-ML techniques. For better reap, exchange off among the schooling pace and the accuracy, an adaptive (ADCNN) is carried out dynamically on the CWRU dataset to trade gaining knowledge. The entire fault analysis version employs a fault pattern dedication thing with the usage of 1-ADCNN and a faulty length assessment element by means of three ADCNN and 3-layer CNN with maximum pooling. The Classification outcomes reveal that the ADCNN has better accurateness when compared to standard CNN and SVM approaches in phrases when figuring out the element rolling disorder. The proposed ADCNN can forecast the error size (illness size) with good precision. On the pinnacle of the conventional structure of CNN, a displaced layer is supplied [33] which could extract the connection among signals with exceptional intervals in paperwork, especially at some point in the exchange of functional situations. it is stated that pleasant accuracy of 96.32% is completed through a divulge phase issue k=three, although the accurateness of conventional CNN without this divulge level is most effective at eighty-three. Like earlier work has a four-layer CNN arrangement with two convolutional and a pair of pooling layers using the CWRU dataset by Qian Peng Company's dataset in China, the correctness overtakes the conventional-SVM and the narrow SoftMax regression classifier, mainly while the vibrating signal is blended with ambient noise. Improvement may be as massive as 25%, showcasing terrific in-built de-noising talents of CNN's set of rules.

A fusion through sensor method is implemented in each of the temporal and spatial data of the CWRU information from 2 accelerometers on the power end, the fan stop is piled by converting 1D time-collection information right into 2D matrix format. Common accurateness by sensors fusion is accelerated to 99.41% from the earlier 98.35% by a best sensor.

Repetitive Neural Network (RNN):

Unique in relation to a feed-forward brain organization, a repetitive brain organization (RNN) processes the information in repetitive conduct, and its engineering is displayed with a stream method which goes from the secret layer to itself that when unfolded in succession, this tends to appear as feed-forward brain system in information grouping. The consecutive model can catch and build successive connections in consecutive information or time-series information. In any case, frequently prepared by back engendering through the period, RNN the infamous angle disappearing/detonating issue originated from its inclination. Albeit RNN is suggested in the 1980s, this has been restricted application because of this explanation, until the introduction, (LSTM-Long Transient Memory) in 1997. “LSTM is expanded with the addition of intermittent entryways named "neglect" doors. Intended to survive the slope disappearing/detonating concern, LSTM shows startling capacity in remembering and presenting drawn-out reliance on information, in this manner playing a predominant job in time-series and printed information examination.

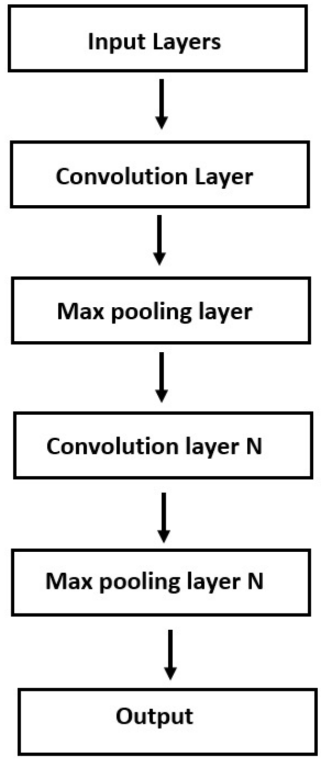

Figure 7: CNN working module

Up to this point, it has gotten incredible triumphs in the field of discourse penmanship acknowledgment, normal language handling, video investigation, and so on. Probably the most primitive usage of RNN was to bear the problem of diagnostics which is accounted for in 2015 here limitations are right off the bat extricated utilizing the discrete wavelet change and later chosen considering the symmetrical fluffy area discriminative examination. These highlights are then taken care of into an RNN for performing bearing issue identification. Test results have revealed that the suggested plot considering RNN is able to precisely identify and order the bearing shortcoming. (RNN-HI) is another RNN-related well-being pointer that is presented for foreseeing RUL of direction by means of LSTM cells utilized in RNN stages. Alongside "time-recurrence" includes, the related correspondence highlight works out likeness among the present observed facts and facts at an underlying working opinion. After playing out connection and monotonicity metrics related including determination procedure, these highlights are relocated to an RNN organization for anticipating the bearing Greetings, through this RUL is assessed. With the information data sets gathered from the originator course of the wind turbines, recommended RNN-HI is exhibited to present improved execution compared to a SOM-related strategy.

GAN: Generative model as Adversarial Network (GAN) a framework was introduced through Goodfellow in 2014. This model proved to be very successful and is playing a vital role in Deep learning. Working as "It depends upon two essential elements such as generator FG which would generate an image from the given data set and discriminator FD which would decide the probability of the produced image is real or fake after being trained. Though a few researchers have worked on the GAN model, it has been proven efficient than most of the other models in fault diagnostics. Plus, GAN is likewise blended with (ADASYN- Adaptive Synthetic Sampling) method for achieving significant oversampling due to the reason, actual testers are very limited. Assessment in opposition to trendy oversampling procedures indicatesthe advantage of embracing GAN. A unique method for fault inquiry related totally to (DCGAN-Deep Convolution GAN) through unjust data sets is suggested. Fresh DCGAN version with four convolutional stages which serve as discriminators and producers is modeled and then implemented on data and extreme/imbalanced signs. The appearing facts synchronize use of the DCGAN version, arithmetical structures built time-domain and frequency-domain facts which are mined to educate the SVM classifier to bear the fault type.

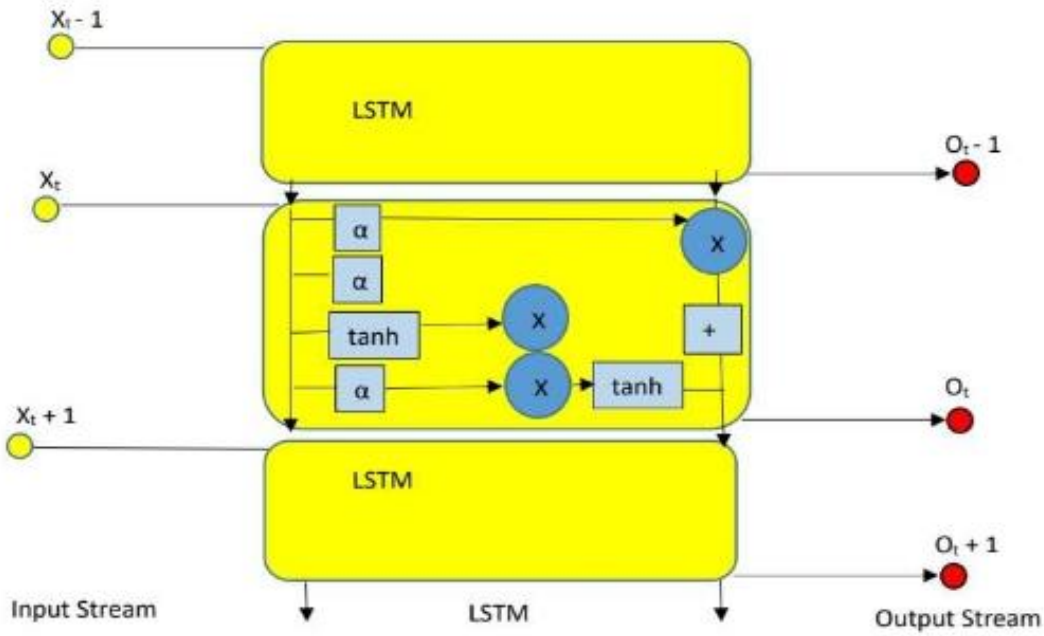

Figure 8: RNN block diagram

DCGAN technique shows higher overall performance than different magnificence balancing methods, such as the "random over-sampling, random underneath sampling, and synthetic minority over-sampling" method. Discovering several studies within the area of bearing fault diagnostics having GAN and related variations for statistics escalation purposes because of their remarkable generative competency.

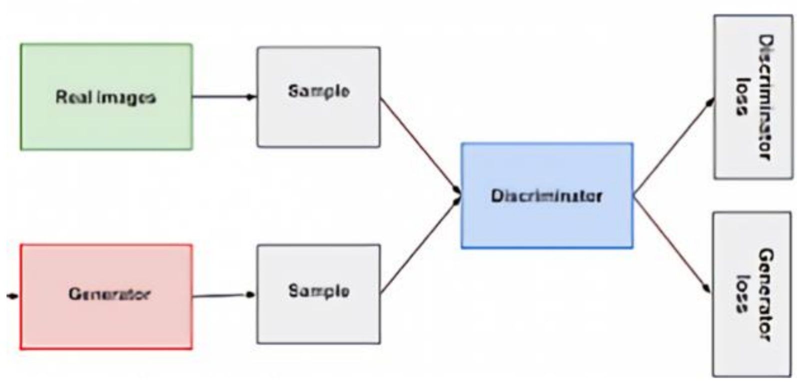

Figure 9 working of GAN[40]

The process during GAN observes knowledge from unlabeled samples unmonitored way, it may moreover analyze the spreading of data in the area which differentiates information that is unidentified classes. In this way, the GAN discriminator is subtle as the classifier aided through a few other components in the structure. The magnificence of GAN-targeted structures is proven superior in semi-supervised fields, where the application is written off as the facts are high-priced and occasional. In another instance, researchers suggested a unique GAN system named explicit (Cat AAE-Adversarial Automobile Encoder), this routinely trains the automobile encoder via an adverse schooling method and enforces a previous circulation at the underlying coding area.

In the subsequent phase, the classifier attempts to group the source input values via the balanced info between samples and their expected categorical magnificence supplies."The latent coding space and the schooling strategies are provided "to examine the gain ofthe suggested version. inspired with the aid of GAN, the latest hostile adaptive-(1D CNN) version-(A2CNN) is recommended to deal with the hassle. Experimentations display shows A2CNN is a robust"fault-discriminative and domain".

Conclusion:

In this review paper, an analysis is drawn to predict the accuracy difference between traditional Machine Learning algorithms later part compares the most popular Deep learning models on the CWRW dataset which gives an accuracy of 99.20% on the GAN model.

Table 2: Accuracy chart

Feature Extraction Algorithm	Number of hidden layers	classifier	characteristics	Accuracy
CNN	4	SoftMax	Noise resilient	92.60%
Deep RNN	3	N/A	accurate	94.75%
GAN	8	SoftMax	Data augmentation	99.20%

References

1. "Report of large motor reliability survey of industrial and commercial installations, Part I," IEEE Trans. Ind. Appl., vol. IA–21, no. 4, pp. 853-864, Jul./Aug. 1985.
2. "Report of large motor reliability survey of industrial and commercial installations, Part II," IEEE Trans. Ind. Appl., vol. IA–22, no. 4, pp. 865–872, Jul./Aug. 1985.
3. K. Teotrakool, M. J. Devaney and L. Eren, "Bearing fault detection in adjustable speed drives via a support vector machine with feature selection using a genetic algorithm," in Proc. IEEE Instrum. Mea. Technol. Conf., Victoria, BC, Canada, 2008, pp. 1129–1133.
4. Soft Comput., vol. 11, no. 6, pp. 4203–4211, Sep. 2011. [51] T. Benkedjouh, K. Medjaher, N. Zerhouni and S. Rechak, "Fault prognostic of bearings by using support vector data description," in Proc. IEEE Conf. Progno. Health Manag., Denver, CO, USA, 2012, pp. 1–7.
5. (PHM), Beijing, China, 2017, pp. 1–5. [57] B. R. Nayana and P. Geethanjali, "Analysis of statistical time-domain features effectiveness in identification of bearing faults from vibration signal," IEEE Sensors J., vol. 17, no. 17, pp. 5618–5625, 1 Sept.1, 2017.
6. M. Cococcioni, B. Lazzerini and S. L. Volpi, "Robust diagnosis of rolling element bearings based on classification techniques," IEEE Trans. Ind. Informat., vol. 9, no. 4, pp. 2256–2263, Nov. 2013.
7. A. Widodo, D. Satrijo, T. Prahasto, G.-M. Lim, and B.-K. Choi, "Confirmation of Thermal Images and Vibration Signals for Intelligent Machine Fault Diagnostics," International Journal of Rotating Machinery, vol. 2012, pp. 1–10, 2012.
8. G.-m. Lim, Y. Ali, and B.-s. Yang, "The Fault Diagnosis and Monitoring of Rotating Machines by Thermography," in Engineering Asset Man- agreement and Infrastructure Sustainability, J. Mathew, L. Ma, A. Tan, M. Weijnen, and J. Lee, Eds. Springer London, 2012, pp. 557–565.
9. C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," Data Mining and Knowledge Discovery, vol. 2, pp. 1–43, 1998.

10. P. Nectoux et al., “Pronostia: An experimental platform for bearings accelerated life test, in Proc. IEEE Int. Conf. Prognostics Health Manag., Denver, CO, USA, 2012, pp. 18.
11. W. Mao, Y. Liu, L. Ding and Y. Li, “Imbalanced fault diagnosis of rolling bearing based on generative adversarial network: A comparative study,” IEEE Access, vol. 7, pp. 9515–9530, 2019
12. S. Suh, H. Lee, J. Jo, P. Lukowicz, and Y.-Oh Lee, “Generative Oversampling Method for Imbalanced Data on Bearing Fault Detection and Diagnosis”, Appl. Sci., vol. 9, no. 4, 746, 2019
13. Liu, H., Motoda H. (1998) Feature Selection for Knowledge Discovery and Data MiningFeature Selection for Knowledge Discovery and Data Mining., Kluwer Academic Publishers. Norwell, MA, USA. 1998
14. "Feature engineering in Machine Learning" (PDF). Zdenek Zabokrtsky. Archived from the original (PDF) on 4 March 2016. Retrieved 12 November 2015
15. S. Zhang, S. Zhang, B. Wang and T. G. Habetler, “Deep Learning Algorithms for Bearing Fault Diagnostics - A Review,” in Proc. IEEE Int. Symp. Diagnostics Electr. Mach., Power Electron. Drives, Toulouse, France, 2019, pp. 257-263.
16. Martin Arjovsky, Amar Shah, and Yoshua Bengio. Unitary evolution recurrent neural networks. In International Conference on Machine Learning, pages 1120–1128, 2016.
17. Liu Buyu. Study on Rolling Bearing Fault Diagnosis Based on Deep Learning. Hangzhou Dianzi University, 2019.
18. M. Cerrada et al., “A review on data-driven fault severity assessment in rolling bearings,” Mech. Syst. Signal Process., vol. 99, pp. 169–196, Jan. 2018.
19. J. Lee, H. Qiu, G. Yu, J. Lin, and Rexnord Technical Services, IMS, University of Cincinnati. “Bearing Data Set” NASA Ames Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA, Accessed: Jul. 2019, [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository.
20. A. Zappone, M. Di Renzo, and M. Debbah, “Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both?,” IEEE Trans. Commun., vol. 67, no. 10, pp. 7331–7376, Oct. 2019.
21. M. Xia, T. Li, L. Xu, L. Liu and C. W. de Silva, “Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks,” IEEE/ASME Trans. Mechatronics, vol. 23, no. 1, pp. 101– 110, Feb. 2018.
22. Z. Zhuang and Q. Wei, “Intelligent fault diagnosis of rolling bearing using one-dimensional multi-scale deep convolutional neural network based health state classification,” in Proc. 2018 IEEE 15th Int. Conf. Netw., Sens. & Control (ICNSC), Zhuhai, China, 2018, pp. 1–6.
23. W. Zhang, C. Li, G. Peng, Y. Chen, Z. Zhang, “A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load,” Mech. Syst. Signal Process., vol. 100, pp. 439–458, 2018.
24. R. Liu, G. Meng, B. Yang, C. Sun and X. Chen, “Dislocated time series convolutional neural architecture: An intelligent fault diagnosis approach for electric machine,”
25. no. 3, pp. 1310–1320, June 2017. [107] C. Lu, Z. Wang, and Bo Zhou, “Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network-based health state classification,” Adv. Eng. Informat., vol. 32, pp. 139–157, 2017.
26. W. Abed, S. Sharma, R. Sutton, and A. Motwani, “A robust bearing fault detection and diagnosis technique for brushless DC motors under non-stationary operating conditions,” J. Control Autom. Electr. Syst., vol. 26, no. 3, pp. 241–254, Jun. 2015.

27. L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin, "A recurrent neural network based health indicator for remaining useful life prediction of bearings," Neurocomput., vol. 240, pp. 98–109, 2017.
28. Y. O. Lee, J. Jo and J. Hwang, "Application of deep neural network and generative adversarial network to industrial maintenance: A case study of induction motor fault detection," in Proc. IEEE Int. Conf. Big Data, Boston, MA, 2017, pp. 3248–3253.
29. A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv:1511.06434, 2015.