Whitepaper

The Power of Generative AI in Predictive Maintenance: Exploring Equipment Failure Cases

Authored by:
Somsuvra Chatterjee, Dr. Anindita Desarkar, and Vishwanathan Raman
LTIMindtree.

Introduction

The advent of GANs has marked a significant turning point in the field of artificial intelligence. GANs have been successfully applied in various domains, ranging from image and video synthesis to text-to-image conversion and anomaly detection. However, one of the most pressing challenges facing the widespread adoption of GANs lies in the quality and diversity of the generated samples. The generative AI market is likely to surpass US$ 167.4 billion by 2033 at a CAGR of 31.3% during the forecast period.[1]

Generative AI includes a set of algorithms that identify patterns within the existing dataset and generate a new set from there while preserving the same pattern. In today's world, its application is quite vast and spread across all domains like healthcare, manufacturing, retail, travel, and hospitality. A few common applications include:

- Image analytics
- Text-to-image synthesis
- Image-to-image translation
- Anomaly detection
- Text-to-speech
- Data augmentation
- Video/music/3D model synthesis

In image synthesis, GAN can generate several images like the original ones, which can be used for model training to enhance accuracy. In anomaly detection, GAN is utilized to identify outliers in data, which can be useful for fraud detection, network intrusion, or medical conditions. Data augmentation can be achieved in a great way by generating several sets of images in various postures. [2]

However, Gen AI consists of three major models: general adversarial networks (GANs), autoregressive convolutional neural networks (AR-CNN), and transformer-based models. This whitepaper focuses on GAN-based models to provide the solution for generating equipment failure datasets. The generic architecture of GAN is provided in Figure 1; it consists of two major components: generator and discriminator.

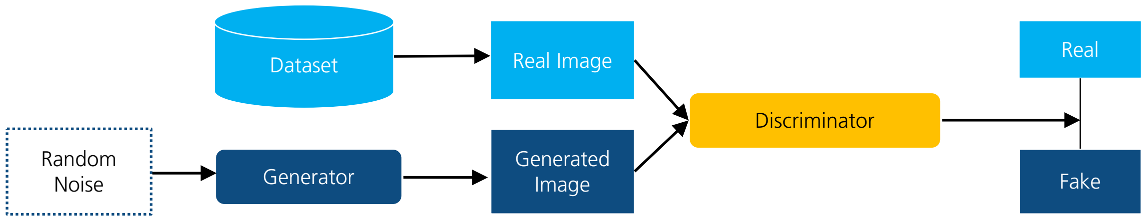


Figure 1: GAN architecture

In the domain of predictive maintenance, we receive sensor datasets from various equipment, but one major problem is the absence of enough real failure samples. It leads to poor accuracy of the failure prediction model. Hence, a GAN-based approach is proposed for generating synthetic datasets for failure cases. [3]

Leveraging generative AI across business functions

According to an MC Kinsey report[5], Figure 2 presents the impact of generative AI on productivity in various business functions, which clearly supports the embracement of the technology across domains.

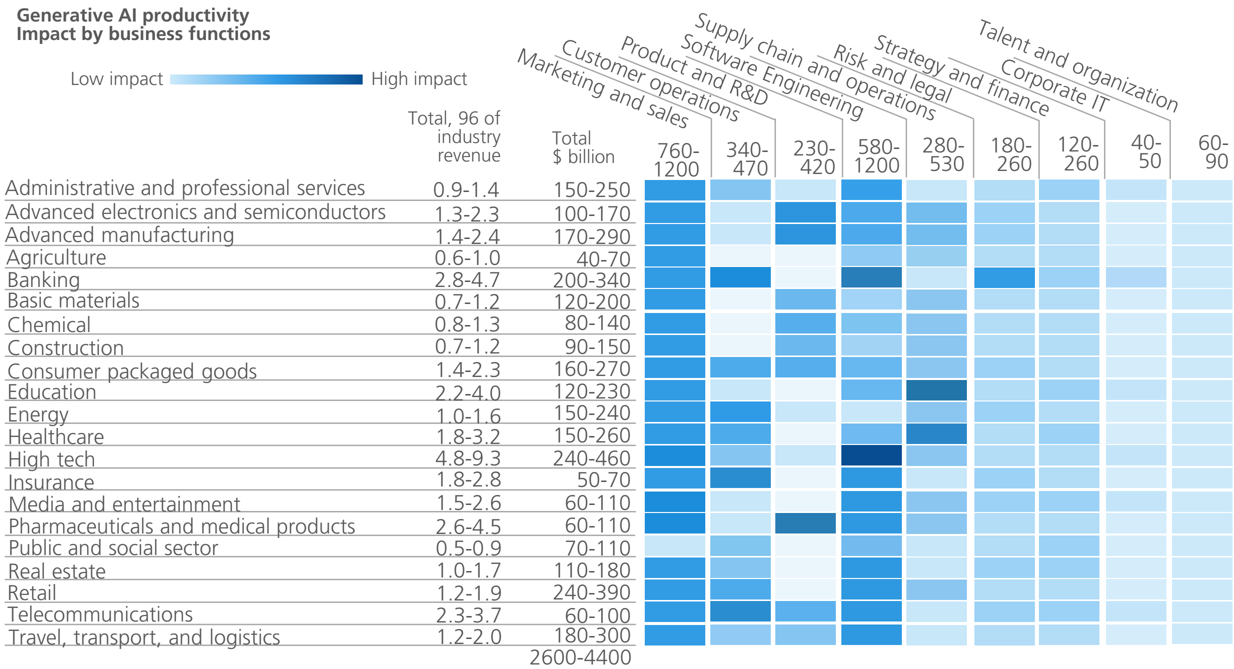


Figure 2: Impact of generative AI on productivity in various business functions [5]

Solution outline

Receiving faulty data for any equipment is always a daunting task. On top of that, problems arise when sufficient records are not available for all kinds of failures.

The present turbine dataset consists of both normal and faulty cases. However, faulty cases are insufficient to get a decent accuracy of the failure prediction model. Faulty refers to cases with abnormal deflection (measured at an amplitude of 150 um) due to high rotational speed in the shaft. Therefore, implementing GAN will generate more anomaly scenarios.

Though the turbine is presented as the equipment of this case study, it can be applied to any other device as the problem is the absence of enough faulty records.

Figure 3 presents the solution outline, which talks about input details, implementing the GAN model towards generating failure cases, and ultimately running the failure prediction model towards receiving better model accuracy. Here, the input is a turbine sensor dataset containing success and failure cases. Sensor values include three different indicator values: time and frequency indicators, waveform indicators, and phase indicators. After processing the input, failure cases are separated and passed through the GAN model to generate more cases. The ultimate objective is to enhance model accuracy and perform better failure prediction .

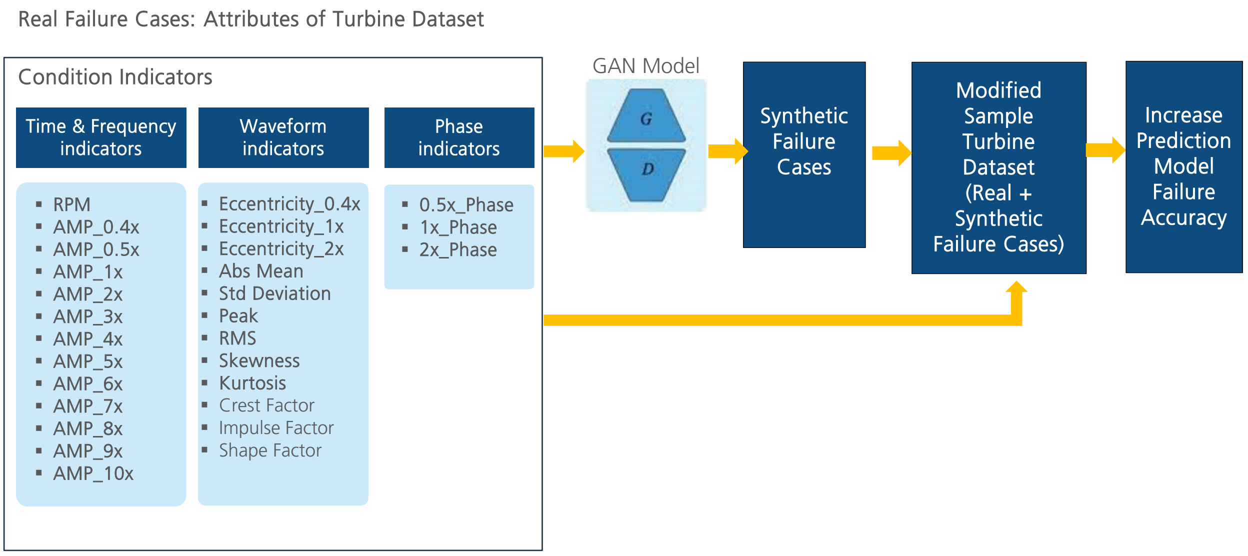


Figure 3: Solution outline

Detailed solution

Figure 4 depicts the steps taken to arrive at the solution.

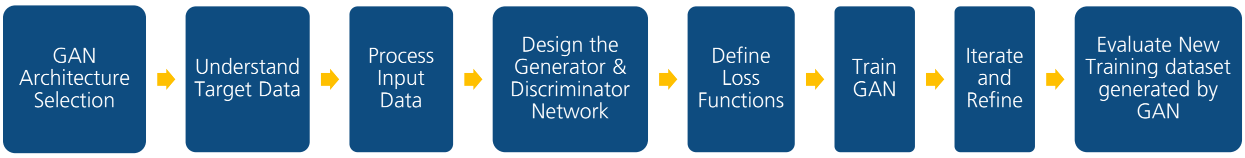


Figure 4: Solution steps

Step 1: Selecting the appropriate GAN model is the first and most crucial task, as getting the most meaningful insight is heavily dependent on that. Here, GAN-FP (GAN for Failure Prediction) is utilized to generate synthetic failure cases.

Step 2: The process design is also dependent on what type of target we want to generate. It's a structured dataset as its asset of failure cases like the original ones.

Step 3: Processing the input dataset is another significant step towards achieving better model accuracy. Hence, appropriate feature engineering is performed here.

Step 4: Designing the generator and discriminator network consists of two major tasks: the generator network receives random noise and feedback from the discriminator, which helps the generator enhance its performance; meanwhile, the discriminator network tries to distinguish between real and fake data. The discriminator gets the input from the training dataset, and the generator's output is combined. Finally, it produces a probability score, which indicates the likelihood of the input being real or fake.

Step 5: GAN utilizes two loss functions: generator loss and discriminator loss. The generator loss determines how efficiently the generator fools the discriminator, whereas the discriminator loss computes the capability of the discriminator to categorize real and generated data.

Step 6: The next or most crucial step is training the GAN model. It predominantly consists of the following steps.

- **Training discriminator:** The discriminator should be trained with a faulty real dataset and the data built by the generator through random noise. These two types of datasets will be tagged with the corresponding label: 0 for generator output and 1 for real ones.
- **Training generator:** It includes passing a bunch of noise data along with true labels to the full GAN model so that the discriminator does not update its weights with these false facts; it indicates that its layers are not trainable.
- Generated noise and the labels are transferred to the GAN to be executed by the generator and the discriminator. It changes the weights of the generator to produce enhanced-quality output at each step. The fake dataset, labeled as real data, is produced by the generator. The discriminator returns to 0 if the data is identified as fake; otherwise, it returns to 1.

Step 7: Refining the target dataset includes alternate training of the discriminator and generator in a series of iterations. As the GAN trains, the generator discovers how to produce more accurate data, while the discriminator develops its competence to distinguish real data from generated data. This process is continued until it reaches a satisfactory performance level.

Step 8: Evaluation of newly generated data is the last step of the process, which determines the quality of GAN output. The fault prediction model is run by the original dataset and the generated dataset coupled with the original dataset, after which prediction accuracy is compared to whether it improved due to an enhanced faulty dataset.

Comprehensive perspective on GAN performance analysis

To evaluate the effectiveness, here is a performance analysis of GAN and a comparative analysis between original and synthetic data. This sample performance analysis of GAN and a comparative analysis between original and synthetic data shows that synthetic data is nearly the same as the original data.

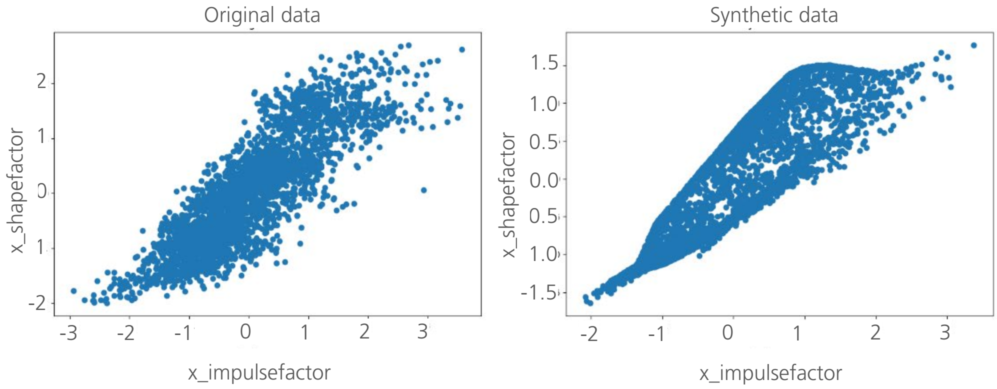


Figure 5: Vanilla GAN outcome: Correlation between two features (x_impulsefactor and x_shapefactor) of the data; original data vs synthetic data

As observed, the correlation pattern between the two datasets is remarkably similar, indicating that the synthetic data accurately captures the underlying relationships present in the original data. This suggests that the synthetic data not only closely resembles the original data but also slightly improves upon it in terms of correlation.

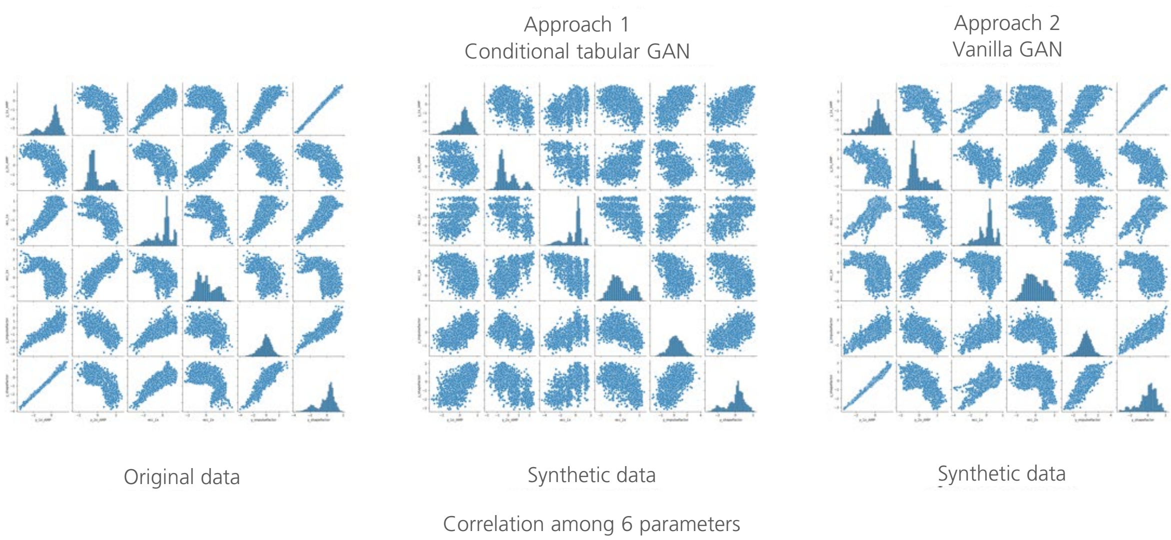
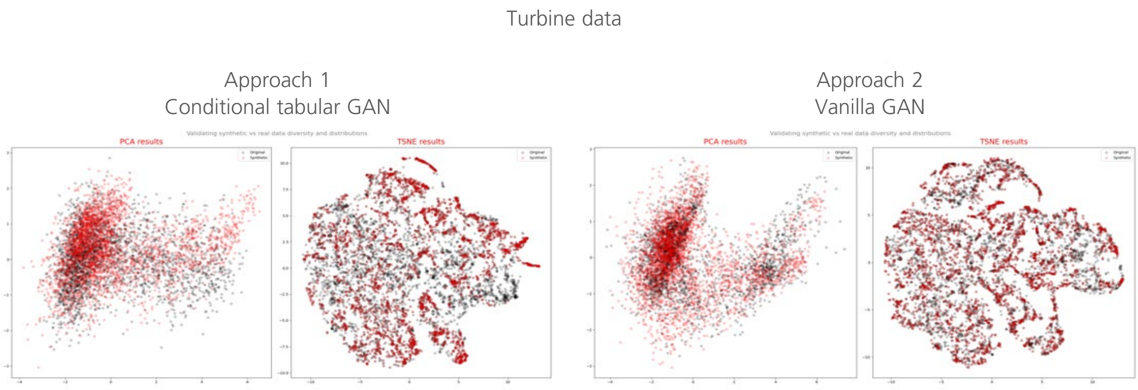


Figure 6: Data validation based on correlation matrix

As seen in the above figure, the original data exhibits a strong correlation between the features, which is preserved in the synthetic data. This indicates that the synthetic data successfully replicates the complex relationships existing in the original data.



Visualization of the data in 2-Dimention based on PCA (two components) and t-distributed stochastic neighbor embedding (TSNE)

Figure 7: Comparison between CT GAN and Vanilla GAN

As evident from the figure, both CT GAN and Vanilla GAN were able to capture the underlying patterns in the original data. However, CT GAN generates more diverse and scattered data compared to Vanilla GAN, which produces data that appears more uniform and less varied. This suggests that CT GAN is better at capturing the complexity and variability present in the original data.

This implies necessity of applications where synthetic data is required to augment or substitute real-world data, such as in machine learning models or simulations.

Unlocking boundless potential Exploring Gen AI: Use cases, ROI, calculations, and research insights

Generative artificial intelligence (Gen AI) has the potential to transform a wide range of industries and domains by enabling the creation of new, unique, and valuable data, products, and services. With its ability to generate novel and diverse outputs, Gen AI can unlock new possibilities for innovation, creativity, and automation. Here are a few cases where Gen AI revolutionizes real-life scenarios in the domain of predictive maintenance. Following Figure 8 presents the same.

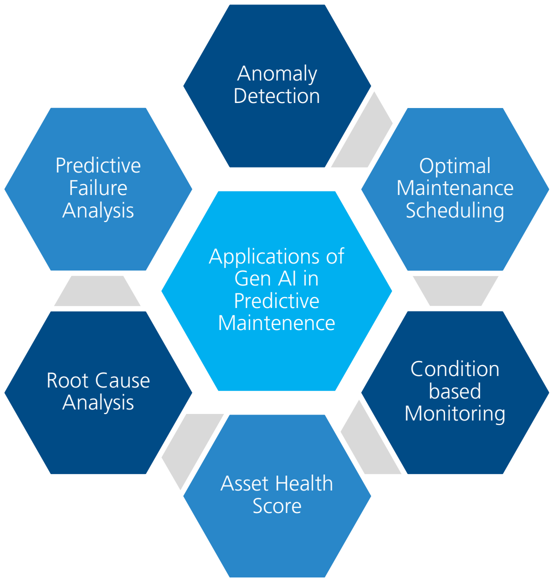


Figure 8: Applications of Generative AI in Predictive Maintenance

Anomaly detection: Initially, the normal, non-anomalous dataset is collected for a machine for which anomaly detection or failure analysis is required to be performed. As a next step, the synthetic dataset is generated based on the normal ones and added to the real dataset so that the model can be trained on the whole set. Anomaly detection comes into the picture when a new dataset is received for the machine. The real data, which significantly deviates from the normal data distribution, is flagged as an anomaly.

By learning the nuisances of normal data, Generative AI can detect anomalies at an early stage, potentially preventing costly issues or saving the machines from catastrophic failure.

Optimal maintenance scheduling: Maintenance scheduling is enhanced using Gen AI by optimizing both employee availability and workload. Several large language models like PaLM 2 can generate optimization solutions from the given objective functions and constraints.

Condition-based monitoring: In condition-based monitoring, appropriate alerts should be generated based on the machines' transmitted data on various components' health and performance. Gen AI models can help in analyzing the data in real-time to assess the condition of the machine.

Asset health score: Generative AI (multimodal models) can integrate diverse data sources (sensor measurements, images and text from manuals, and inspection notes) and generate a more comprehensive equipment health score by understanding its inner details.

Root cause analysis: Gen AI models do the incident analysis more accurately and suggest actions in real-time to solve them. The major activities of automated incident analysis consist of detecting IT incidents, gathering required data, arranging alerts chronologically, identifying root causes, and recommending solutions.

Generative AI in predictive maintenance

Exploration shows how the implementation of Gen AI has maximized the efficiency of industrial operations in the predictive maintenance domain.

Data generation and augmentation

The most common case is the creation of novel failure scenarios and patterns that do not exist in the real dataset so that models can be trained accordingly on these potential failures. Business benefits include early detection of anomalies, reduced downtime, cost savings, and improved decision-making. Predictive models become more powerful as innovative failure cases are included in the training dataset, which leads to better prediction and improved accuracy toward predicting outlier scenarios in advance. As a result, equipment downtime is also reduced as appropriate measures can be taken in advance. This leads to huge cost savings as production is not hampered due to early prediction. Also, Generative AI provides valuable insights into equipment behavior and performance, enabling data-driven decision-making for maintenance strategies, spare parts management, and resource allocation.

Conditional GAN: generate synthetic signals for digital twin

Generative Adversarial Networks (GANs) can be utilized to create artificial data that closely mirrors genuine input data fed into the networks. This proves advantageous when computational expenses or experimental costs prohibitively restrict the acquisition of real data. Moreover, Conditional GANs (CGANs) can leverage data labels during training to fabricate data conforming to predetermined categories.

3D object generation: digital twin

In their 2016 paper, "Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modelling," Jiajun Wu et al. present a Generative Adversarial Network (GAN) capable of generating novel three-dimensional objects, including examples such as chairs, cars, sofas, and tables. This achievement marks a significant milestone in the field of 3D generative modeling [8]. The 2016 paper titled "3D Shape Induction from 2D Views of Multiple Objects," (by Matheus Gadelha, Subhransu Maji, Rui Wang) uses GANs to generate three-dimensional models given two-dimensional pictures of objects from multiple perspectives.

Anomaly detection

Generative models can be trained on normal datasets to understand equipment behavior in good condition. More samples can be generated by following the current pattern. Any deviation from normal behavior can be identified as an anomaly.

Failure mode simulation

Getting equipment faulty data is always difficult, especially in all the scenarios. Gen AI models can simulate other error situations for which real data might not be available. It will enhance the accuracy of the failure detection model.

Reducing data imbalance

In the real scenario, data imbalance is a major problem while training the prediction model because faulty data is much less compared to the normal dataset. As an example, it will be very small in number if the equipment is newly installed. In those cases, Gen AI models can generate more faulty cases by following the pattern of the available dataset.

Optimize maintenance schedule

By implementing Gen AI models, we can generate synthetic data for failure cases and can simulate various abnormal situations. It leads to the optimization of equipment maintenance schedules as failure cases are predicted better.

Cost minimization

Total operational cost is minimized as failure cases are predicted in a better way as it leads to minimum downtime of equipment and avoids production loss.

Conclusion

This study presents a novel GAN-based failure case generation framework that helps enhance the failure model prediction accuracy. Various experiments on the generated dataset clearly show the similarity with the original dataset. This will revolutionize the domain of predictive maintenance as data scarcity is a major concern for most equipment, especially faulty cases. Future work considers applying the same framework to other equipment datasets to understand the overall performance. Additionally, more GAN-based models can be tried to check for the most appropriate one.

References

1. *Generative AI Market Outlook (2023 to 2033),* FMI: https://www.futuremarketinsights.com/reports/generative-ai-market

2. *A Generative Adversarial Network Based a Rolling Bearing Data Generation Method Towards Fault Diagnosis,* Lin Huo | Huanchao Qi | Simiao Fei | Cong Guan | Ji Li, Hindawi, 13 Jul 2022: https://www.hindawi.com/journals/cin/2022/7592258/

3. *An Introduction to Image Synthesis with Generative Adversarial Nets,* He Huang | Philip S. Yu | Changhu Wang, arxiv, 17 Nov 2018: https://arxiv.org/abs/1803.04469

4. *Generative Adversarial Network-Based Fault Detection in Semiconductor Equipment with Class-Imbalanced Data*, Jeong Eun Choi | Da Hoon Seol | Chan Young Kim | Sang Jeen Hong, MDPI, 8 Feb 2023: https://www.mdpi.com/1424-8220/23/4/1889

5. *The economic potential of generative AI: The next productivity frontier,* McKinsey, Report, 14 Jun 2023: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

6. *Intelligent Fault Diagnosis of Rotary Machines: Conditional Auxiliary Classifier GAN Coupled With Meta Learning Using Limited Data,* Sonal Dixit | Nishchal K, IEEE, 20 May 2021: https://ieeexplore.ieee.org/document/9437218

7. *A Modified Generative Adversarial Network for Fault Diagnosis in High-Speed Train Components with Imbalanced and Heterogeneous Monitoring Data,* Chong Wang | Jie Liu | Enrico Zio, JDMD, 02 Apr 2022: https://ojs.istp-press.com/dmd/article/view/68

8. *Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling,* Jiajun Wu | Chengkai Zhang | Tianfan Xue | William T. Freeman | Joshua B. Tenenbaum, arxiv, 04 Jan 2017: https://arxiv.org/abs/1610.07584

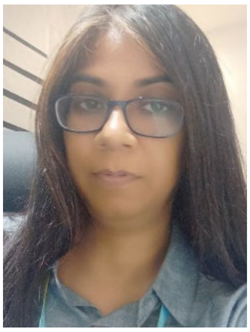
9. *3D Shape Induction from 2D Views of Multiple Objects,* Matheus Gadelha | Subhransu Maji | Rui Wang, IEEE, 18 Dec 2016: https://arxiv.org/abs/1612.05872

Author bio

Somsuvra Chatterjee

Director – Enterprise AI, LTIMindtree

Somsuvra Chatterjee is a committed and passionate data scientist with 15+ years of experience, of which 12+ years were in the field of analytics and data science. He has worked across diversified domains, including CPG, Retail & Manufacturing, Energy, etc. He was engaged in developing solutions for various computer vision and image processing-based products and projects.



Dr. Anindita Desarkar

Associate Director – Enterprise AI, LTIMindtree

Anindita Desarkar is a dedicated researcher and IT professional with 19+ years of experience in the IT industry. She has 18+ publications in Scopus-indexed journals and conferences. Aninidita completed her PhD thesis on "Machine Learning and Analytics" from the Department of Computer Science & Engineering, Jadavpur University. She was involved in developing solutions for various machine learning and optimization-based products and projects.

Vishwanathan Raman

Principal Director – Enterprise AI, LTIMindtree

Vishwanathan Raman is a technology leader, a technical author, an AI strategist, a storyteller, and a programmer at heart with over 26+ years of experience across the globe. He leads the AI Competency Development and Generative AI Research at LTIMindtree.