

Research Progress on Big Data Feature Mining of Electromagnetic Induction Effects in Electric Motors

Bohan Zhang

Grade 12, Xiaoshi High School,
Ningbo, Zhejiang, China
zbh13065686603@gmail.com

Abstract:

With the rapid development of Industry 4.0 and intelligent manufacturing, electromagnetic induction effects in motors, such as eddy current effects, harmonic distortion, and hysteresis losses, have become key factors affecting the energy efficiency and lifespan of motors. However, traditional analytical methods are highly dependent on limited working condition simulations and experimental data. This paper systematically reviews the application progress of big data technology in the study of electromagnetic induction effects in motors, with a focus on analyzing the advantages and disadvantages of traditional simulation, machine learning, and deep learning methods in feature extraction, anomaly detection, and optimization design. Research shows that data-driven technologies have significantly enhanced the accuracy and efficiency of energy efficiency optimization, fault diagnosis, and intelligent control. Future research needs to further integrate multi-source data with physical models to provide theoretical support and practical guidance for intelligent design, predictive maintenance, and energy efficiency optimization.

Keywords: Electromagnetic induction effect; big data feature extraction; anomaly detection; intelligent design.

1. Introduction

In the modern industry and energy systems, electric motors are core power equipment and are widely used in many fields such as electricity, transportation, construction, and manufacturing. However, with the increase in motor power and speed, electromagnetic induction effects such as eddy current effect and harmonic distortion, have gradually become key factors

that affect motor energy efficiency and reliability. Eddy current effect will generate additional heat loss during motor operation to reduce motor efficiency, while harmonic distortion may cause motor operation instability and even cause equipment failure. Therefore, how to effectively identify and control these electromagnetic induction effects is one of the main issues in improving motor performance and energy efficiency. Traditional motor analysis methods main-

ly rely on small sample data and empirical models, which are difficult to completely reflect electromagnetic behavior under complex working conditions. For instance, experimental data acquisition costs are high, the cycle is long, and it is difficult to cover all potential operating states. And simulation models which is based on limited data often have limited accuracy and are difficult to meet the needs of modern industry for high precision and accurate analysis. Additionally, when dealing with large-scale motor systems, traditional methods lack the ability to identify abnormal signals which will result in delayed fault diagnosis and affecting equipment maintenance efficiency. In recent years, the rapid development of big data technology has provided new opportunities for the analysis and optimization of motor electromagnetic induction effects. Industrial Internet of Things technology enables real-time monitoring and data collection of motor operating status, thereby building a large-scale, multi-dimensional operating data set. Based on this data, machine learning, deep learning and other algorithms can be used to automatically extract and analyze electromagnetic induction features to achieve anomaly detection and fault warning. For example, by analyzing the current, voltage, temperature and other parameters of the motor during operation, potential eddy current effects or harmonic distortion problems can be identified, and optimization measures can be taken in advance, such as adjusting the motor design, optimizing the control strategy. Therefore, it can improve the overall energy efficiency and reliability of the motor to some extent.

This research will focus on the application progress of big data technology in the electromagnetic induction feature extraction, anomaly detection and optimization design of motors. And the scale of this search will cover three aspects: data acquisition, feature mining and application implementation, which aim to compare the advantages and disadvantages of different technical routes and explore their practical value in intelligent motor design, predictive maintenance and energy efficiency optimization. It provides theoretical support and practical guidelines for the intelligent and green development of future motor systems through comprehensive analysis of the combination of big data and electromagnetic induction technology.

2. Big Data Sources and Preprocessing of Electromagnetic Induction Effects

The research on electromagnetic induction includes many aspects, such as the temporal and spatial distribution of electromagnetic fields, signal propagation characteristics, and noise interference identification. Its data sources are

extensive, including simulation data, operation detection data, and environmental data. With the development of big data technology, how to efficiently obtain and process this data is an important prerequisite for promoting the intelligent research of electromagnetic induction effect.

2.1 Data Source and Type

The data sources of electromagnetic induction effect mainly include the following categories:

Simulation data: high precision electromagnetic field data generated by electromagnetic simulation software such as COMSOL, ANSYS, to simulate the electromagnetic induction process. Simulation data usually includes information such as electromagnetic field strength, phase distribution, frequency response and so on, which can reflect the theoretical characteristics of electromagnetic induction.

Operation detection data: Real-time monitoring data from actual operating equipment (such as motors and generators), including signals like current, voltage, temperature and vibration. These data reflect the performance of electromagnetic induction in actual environments and serve as an important empirical basis for studying the effect of electromagnetic induction. For instance, by deploying intelligent sensors, real-time energy consumption data during the operation of equipment can be collected, including multi-dimensional information such as voltage, current, and temperature.

Environmental data: it includes environmental factors such as geographic information, climate conditions, and material properties. These data have a vital influence on the spatiotemporal distribution and signal propagation characteristics of electromagnetic induction. For example, terrain, soil type, humidity will affect the propagation and induction effect of electromagnetic waves.

2.2 Data Preprocessing Process

After obtaining the electromagnetic induction effect data, a series of preprocessing steps are required to improve the data quality and provide a reliable foundation for subsequent analysis.

Noise filtering: Electromagnetic induction data is usually interfered by noise, such as environmental noise, equipment noise. To improve the signal-to-noise ratio of the data, filtering techniques such as low-pass filtering, wavelet transform, can be used to remove noise. What's more, machine learning methods such as wavelet packet decomposition, wavelet threshold denoising, can be used to process the noise more finely.

Data alignment: Since electromagnetic induction data may come from multiple sensors or different time points,

the data needs to be time-aligned to ensure the consistency of the data in the time dimension. For example, in a multi-sensor collaborative monitoring system, data from different sources can be aligned to the same time axis by interpolation or timestamp matching.

2.3 Open-Source Databases and Tools

In order to support big data research on electromagnetic induction effects, some open-source databases and tools are widely used, providing researchers with rich data resources and analysis tools.

NASA Motor Degradation Dataset: This dataset contains the degradation process data of the motor running state, including parameters such as current, voltage, and temperature. These data can be used to study the fault evolution process of the motor and the degradation mechanism of the electromagnetic induction effect.

IEEE PHM Challenge Dataset: This dataset is mainly used for predictive maintenance research and contains multi-source data of the equipment running state, such as vibration, temperature, and current. These data can be used to study the performance of the electromagnetic induction effect in the equipment degradation process.

TensorFlow for feature learning analysis tools: TensorFlow is a powerful open-source machine learning framework that supports deep learning models such as convolutional neural networks and recurrent neural networks and is widely used in signal feature extraction. In the study of electromagnetic induction effects, TensorFlow can be used to build deep learning models (such as 1D CNN, LSTM) to automatically extract the spatio-temporal features of electromagnetic signals (such as current and voltage waveforms), thereby enhancing the level of automation in feature extraction.

Big data research on electromagnetic induction effects relies on diverse data sources and efficient preprocessing methods. Simulation data, operation test data, and environmental data form the basis of electromagnetic induction effect research, while preprocessing steps such as noise filtering and data alignment ensure the quality and consistency of the data. Open-source databases and tools such as NASA motor degradation dataset, IEEE PHM Challenge dataset, and TensorFlow, provide rich data resources and powerful analysis tools for electromagnetic induction effect research, promoting the intelligent development of research.

3. Big Data Feature Mining Methods for Electromagnetic Induction Effects

With the rapid development of big data and artificial in-

telligence technology, the research on electromagnetic induction effect has gradually developed from traditional experimental analysis to data-driven intelligent analysis. The electromagnetic induction effect involves many aspects such as the spatiotemporal distribution of electromagnetic fields, signal propagation characteristics, and noise interference identification. Its data characteristics are complex, so it is necessary to use advanced analysis methods for mining and modeling. The content below will systematically explain the big data characteristics for mining methods of electromagnetic induction effect from three different aspects: traditional statistical and signal analysis methods, machine learning driven methods, and deep learning driven methods.

3.1 Traditional Statistical and Signal Analysis Methods

Traditional statistical and signal analysis methods have significant application value in the big data feature mining of electromagnetic induction effects of motors. These methods can extract meaningful features by analyzing signals such as current, voltage, and temperature of motors to support subsequent analysis and decision-making. They mainly include time-domain and frequency-domain analysis, Hilbert-Huang transform (HHT), and correlation analysis, etc.

Time-domain analysis is one of the traditional signal processing methods. It directly processes the information of signal changes over time and can intuitively reflect the dynamic characteristics of the signal. In the research of electromagnetic induction effects in motors, time-domain analysis is often used to identify transient phenomena and periodic changes in signals, thereby providing a basis for fault diagnosis and condition monitoring. For instance, in the fault diagnosis of induction motors, by analyzing the time-domain characteristics of current signals, such as root mean square (RMS), standard deviation, peak factor, kurtosis and skewness, etc., bearing faults can be effectively identified [1]. In addition, time-domain analysis can also be used for the analysis of vibration signals. By observing the waveform characteristics of the signals, abnormal states of mechanical systems can be identified [2]. Frequency-domain analysis is to convert a signal from the time domain to the frequency domain in order to analyze the frequency components of the signal. This method is particularly useful in the analysis of electromagnetic induction effects in motors, as it can help identify harmonic components in the signal, thereby evaluating the operating status of the motor. Commonly used frequency-domain analysis methods include Fourier Transform (FT), Short-time Fourier Transform (STFT), etc [3].

The Hilbert-Huang Transform (HHT) is a method that combines the Hilbert transform and empirical mode decomposition. HHT overcomes the defect of traditional time-frequency analysis methods that represent non-stationary signals with meaningless harmonic components and can achieve extremely high time-frequency resolution and has good time-frequency aggregation [4].

Correlation analysis is used to evaluate the statistical relationship between different signals or variables. During the operation of a motor, due to the existence of current harmonics, the magnetic field distribution will be uneven, thereby generating additional eddy current losses in different components of the motor (such as stator teeth, yoke parts, rotor, etc.). These eddy current losses will further cause local temperature to rise, forming hot spots. Therefore, by analyzing the correlation between the readings of temperature sensors at different positions and the harmonic content of three-phase currents, the hotspots of eddy current losses can be effectively identified [5].

3.2 Machine Learning Driven Methods

With the development of machine learning technology, the study of electromagnetic induction effects has gradually introduced supervised learning and unsupervised learning methods in the sake of improving the efficiency of feature extraction and pattern recognition.

Supervised learning trains models by using labeled data, enabling them to classify or predict new input data. In the research of electromagnetic induction effects, supervised learning can be applied to signal classification, noise identification and anomaly detection. For instance, by training a model to recognize electromagnetic signals under normal conditions, the model can automatically detect abnormal signals. This supervised learning method has achieved good results in anomaly detection in power systems and can effectively identify abnormal electromagnetic signals. Unsupervised learning uses clustering, dimensionality reduction and other methods to mine the intrinsic structure of data without labeled data. In the study of electromagnetic induction effects, unsupervised learning can be used for feature extraction and data clustering. For example, methods based on principal component analysis or autoencoders can be used for dimensionality reduction to extract the main features or characteristics of electromagnetic signals. As a result, it can reduce the dimension of the data and improving computational efficiency [6]. In addition, unsupervised learning can also be used to identify potential patterns in electromagnetic signals, such as the distribution distinctions of different frequency components, to provide a basis for subsequent signal processing and analysis [7].

3.3 Deep Learning-Driven Methods

Deep learning is an important branch of machine learning which can automatically extract high-level features of data through multi-layer nonlinear transformations and is widely used in the study of electromagnetic induction effects.

In the research of the electromagnetic induction effect of electrical machinery, temporal feature extraction is one of the key links. During the process of time series feature extraction, deep learning has significant advantages because it can automatically extract key features in time series. For example, methods based on long short-term memory networks (LSTMs) can be used to extract time features of electromagnetic signals, such as periodicity, trend, and mutation. LSTM effectively captures long-term dependencies in time series through a gating mechanism, so that improving the accuracy of signal recognition and classification. In addition, methods based on convolutional neural networks (CNNs) can be used to extract local features of electromagnetic signals, such as frequency components and amplitude changes, to provide support for signal classification and recognition. As for the Spatial feature extraction, it mainly focuses on the distribution characteristics of electromagnetic signals in the spatial dimension. It can significantly improve the capability of modeling the spatial futures by multi-scale feature extraction and attention mechanisms. For example, deep learning methods that is based on attention mechanisms can be used to extract spatial features of electromagnetic signals, such as electromagnetic field strength and polarization direction at different locations. By introducing channel attention mechanisms and spatial attention mechanisms, the modeling ability of representing key features can be enhanced and the accuracy of signal recognition can be improved as well [8].

The big data feature mining methods for electromagnetic induction effects cover traditional statistical and signal analysis methods, machine learning driven methods, and deep learning driven methods. Traditional methods play a vital role in basic research, while machine learning and deep learning methods provide more powerful tools for the study of electromagnetic induction effects. With the great development of deep learning technology, its application in the study of electromagnetic induction effects will be more extensive and consequentially provide new solutions for modeling the spatiotemporal distribution of electromagnetic fields, signal recognition and prediction. Future research can further explore the combination of deep learning and physical prior knowledge to improve the generalization ability and robustness of the model and promote the intelligent development of electromagnetic

induction effect research.

4. Application Scenarios of Big Data Feature Mining

With the development of big data technology, its application in many fields is becoming more and more extensive. Particularly, it is widely used in the energy efficiency optimization, fault diagnosis and prediction, and intelligent control, big data feature mining technology. The following will elaborate on its application scenarios from these three aspects.

4.1 Energy Efficiency Optimization--Wind Turbine Eddy Current Loss Optimization, Permanent Magnet Motor Topology Optimization

In the field of wind power generation, energy efficiency optimization is the key to improving energy utilization efficiency. Big data feature mining technology can identify important factors that affect energy efficiency, such as eddy current loss and permanent magnet motor topology, through in-depth analysis of wind turbine operation data. By extracting key features during the operation of wind turbines, such as current, voltage, temperature, an efficient energy efficiency model can be established to optimize the motor design and operation strategy. For example, feature extraction methods that is based on deep learning can automatically identify the main sources of eddy current loss and propose optimization solutions to reduce energy loss and improve power generation efficiency [9]. Also, the topology optimization of permanent magnet motors can also achieve optimal design through big data analysis, plus combination with historical operation data and simulation results, thereby improving overall energy efficiency.

4.2 Fault Diagnosis and Prediction

Fault diagnosis and prediction is one of the important applications of big data feature mining technology. By analyzing the equipment operation data, potential fault modes can be identified and the occurrence time of the fault can be predicted, hence it can achieve early warning and maintenance. In the aspect of early warning of bearing faults, big data technology can extract features such as vibration signals and temperature changes to establish a fault diagnosis model in order to achieve real-time monitoring and prediction of bearing status. For example, feature extraction methods based on machine learning can identify early characteristics of bearing faults, such as abnormal vibration and temperature rise, thereby providing a basis for maintenance [10]. In the aspects of life prediction of motor insulation degradation, big data technology

can analyze the performance data of insulation materials, such as temperature, humidity, aging degree. It can contribute to establish a life prediction model and predict the service life of insulation materials and provide a scientific basis for equipment maintenance.

4.3 Intelligent Control

Intelligent control is another important application area of big data mining technology. By analyzing the system operation data, the control strategy can be optimized and the automation level of the system can be improved. In motor control, big data mining technology can identify key control parameters and optimize the control strategy by real-time monitoring and analyzing the system operation status. In the actual industrial production process, many components and electrical equipment are involved. Especially when these devices are constantly in operation, many parts and equipment are prone to wear and tear faults. Once these faults occur, they will lead to a decline in the overall project efficiency and cause varying degrees of economic losses to the enterprise. Applying the big data intelligent control system to the process of industrial electrical automation construction can better reflect the problems and deficiencies that occur in the assembly line. With the support of electrical automation technology, intelligent operation can be truly achieved and various fault problems can be identified [11].

In summary, big data feature mining technology has broad application prospects in energy efficiency optimization, fault diagnosis and prediction, and intelligent control. Through in-depth analysis and mining of data, comprehensive monitoring and optimization of equipment operating status can be achieved, and the operating efficiency and reliability of the system can be improved.

5. Conclusion

This research systematically explores the application of big data feature mining in electromagnetic induction effects and motor control. Compared with traditional methods, deep learning has achieved outstanding automation in extracting complex spatio and temporal patterns from electromagnetic signals. And it eliminates the reliance on manual feature engineering. More importantly, mechanization and deep learning have greatly enhanced the diagnostic level which enable highly accurate detection and classification of early subtle fault features, such as specific harmonic distortion and minor eddy current abnormalities. However, even so, the scarcity of data remains a significant obstacle to research. Effective fusion methods of multimodal data such as electromagnetic, thermal, mechanical and material are also of vital importance. In addition, the

data-driven approach has established a new paradigm for motor design optimization, effectively handling complex electromagnetic-thermal-mechanical couplings to achieve multi-objective topology refinement. Future research can further integrate physical models with big data analysis to enhance the generalization ability and robustness of the models. And Maxwell's equations can also be explicitly embedded to enhance the model's extrapolation and interpretability. The optimization model for resource-constrained edge devices will achieve real-time feature extraction and adaptive control, while explainable artificial intelligence (XAI) technology is crucial for establishing trust through explainable fault attribution in electromagnetic feature recognition.

References

- [1] Li Junqing et al. Motor Bearing Fault Diagnosis Method Based on current signals and deep reinforcement learning, *Electric Power Science and Engineering*, 2003, 3, (39), pp. 61-70
- [2] Hyeon Bae, Sungshin Kim, Yon Tae Kim, et al. Application of Time-Series Data Mining for Fault Diagnosis of Induction Motors, *Journal of Mechanical Science and Technology*, volume 19, issue 6, pp. 1770-1778
- [3] R. Leidhold, D. Kalugin, O. Tolochko, Experimental research of the magnetization and demagnetization process of the vector-controlled induction motor, *IEEE Transactions on Energy Conversion*, early access, May 2024, doi: 10.1109/TEC.2024.3385437
- [4] Shi Jie, Wu Xing, Liu Tao. Diagnosis of Bearing Recheck Faults using HHT Algorithm and Convolutional Neural Network. *Transactions of the Chinese Society of Agricultural Engineering*, 2020, 2, vol. 36, No. 4
- [5] Li Weili, Zhang Xiaolei, Wang Yuhang. Temperature rise characteristics of rotor with sheath permanent magnet servo motor under current-time harmonic action. *Proceedings of the CSEE*, 2023, 43*(12), 4567-4580
- [6] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. *Nature*, 1986, 323(6088): 533-536.
- [7] Wang Manxi, Lu Keyu, Wang Jiabao, et al. Small Sample Recognition Technology Based on Data Augmentation and Ensemble Learning [J]. *Mobile Communications*, 2022, 46(7): 02-09.
- [8] WANG Xiang, DENG Wen, LIU Shixiong, HUANG Zhitao. Anomaly detection method of electromagnetic time series based on attention mechanism[J]. *Journal of Terahertz Science and Electronic Information Technology*, 2021, 19(4): 581-588
- [9] Yongji Zhang. "Electrical Engineering and Intelligent Control Strategies in Renewable Energy Integration: Enhancing Energy Sustainability." *Journal of Theory and Practice of Management Science*, 2024. 04(05). 06, issn: 2790-1491
- [10] D. Gheorghe, Mihai Ioan. "ASPECTS REGARDING DATA MINING APPLIED TO FAULT DETECTION." *Annals of the Oradea University: Fascicle Management and Technological Engineering*, Volume IX(XIX), 2010, NR2
- [11] Shanghai Electric Corporation Limited, Predictive maintenance of intelligent high-end equipment in industrial Internet. <http://www.yrdcpn.com/upload/2023/0822/89710010-6c9d-46ef-917d-5e6135fccfe6.pdf>