Global Energy Interconnection
Development and Cooperation Organization

KeAi
CHINESE ROOTS
GLOBAL IMPACT

Volume 3 Number 2 April 2020 (175-185)
DOI: 10.1016/j.gloei.2020.05.009

Global Energy Interconnection

Contents lists available at ScienceDirect
https: //www.sciencedirect.com/journal/global-energy-interconnection

Full-length article

Research and application of artificial intelligence service platform for the power field

Peng Liu[1], Wei Jiang[2], Xiaohui Wang[1], Hongmin Li[1], Hongjian Sun[3]

1. China Electric Power Research Institute Co. Ltd., Beijing, 100192, P.R. China
2. State Grid Corporation of China, No. 86, West Chang'an Street, Xicheng District, Beijing, 100031, P.R. China
3. University of Durham, The Palatine Centre, Stockton Road, Durham, DH1 3LE, U.K.

Scan for more details

Abstract: Conventional analysis methods cannot fully meet the business needs of power grids. At present, several artificial intelligence (AI) projects in a single business field are competing with each other, and the interfaces between the systems lack unified specifications. Therefore, it is imperative to establish a comprehensive service platform. In this paper, an AI platform framework for power fields is proposed; it adopts the deep learning technology to support natural language processing and computer vision services. On one hand, it can provide an algorithm, a model, and service support for power-enterprise applications, and on the other hand, it can provide a large number of heterogeneous data processing, algorithm libraries, intelligent services, model managements, typical application scenarios, and other services for different levels of business personnel. The establishment of the platform framework could break data barrier, improve portability of technology, avoid the investment waste caused by repeated constructions, and lay the foundation for the construction of "platform + application + service" ecological chain.

Keywords: Artificial intelligence platform, Deep learning, Neural network, Model training, Application scenarios.

1 Introduction

Artificial Intelligence (AI) is currently one of the most subversive technologies, with strong processing power in computational intelligence, perceptual intelligence, and cognitive intelligence. As new technologies are most likely leading toward the fourth industrial revolution, AI is the core support technology of intelligent energy, with strong optimization and learning abilities that deal with high-dimensional, time-changing, and nonlinear problems, which can effectively solve the challenges faced by energy systems.

With the continuous deepening and development of integrated business applications in the energy field, the conventional analytical methods are unable to fully meet the needs of business development, e.g., grid complex network, equipment wide points, and different operating characteristics. In addition, the conventional operation and maintenance methods face difficulty in accurately evaluating the state of the equipment and targeted investment, renovation, and operation. Large-grid wide-area interconnection forms a complex network, the fault identification, optimization

Received: 15 December 2019/ Accepted: 17 February 2020/ Published: 25 April 2020

Peng Liu
liupeng3@epri.sgcc.com.cn

Wei Jiang
wei-jiang@sgcc.com.cn

Xiaohui Wang
wangxiaohui@epri.sgcc.com.cn

Hongmin Li
lihongmin1213@126.com

Hongjian Sun
hongjian.sun@durham.ac.uk

control, and other problems of which are difficult to effectively solve by using existing methods.

In recent years, power systems have attracted considerable interest in the direction of AI research and development, mainly with respect to business systems focusing on the promotion of a certain AI technology. In terms of load prediction, some studied focused on the deep learning algorithm to realize grid-load prediction [2-4], and put forward the nonintrusive load-monitoring model based on the sequence-to-sequence method and attention mechanism to realize the nonintrusive load decomposition application [5]. In terms of fault prediction and processing, Wang [6] and He et al. [7] applied the deep learning algorithm to transformer fault diagnosis. Xu et al. [8] proposed a method to self-learn the fault data and realize fault-type identification of transmission line based on the deep learning framework. Shan et al. [9] proposed a technique of scheduling a fault-management robot based on deep learning. From the aspect of cruise image processing, some studies had adopted the deep learning algorithm to realize the classification of transmission line inspection of aerial photography data [10-13]. Moreover, Cheng et al. [14] used the convolution neural network algorithm to train the sample library rendered by the 3D MAX software to realize the classification of images with an insulator. In terms of transient stability assessment, Hu et al. [15] and Zhu et al. [16] proposed transient stability assessment methods for power systems based on a deep confidence network. By using the deep learning framework, a method of transient stabilization of a power system based on stacked auto-encoders (SAE) was proposed in [17]. In [18], based on the short-term disturbed track after fault removal, a convolution neural network was used in a deep learning model to extract features and construct the nonlinear mapping relationship between the short-term disturbed track and stable category. In terms of safety attack and defense, Wu et al. [19] applied AI technology to maintain the steady-state safety of power grid systems. Wang et al. [20] proposed a defense mechanism based on interval state estimation to detect cyberattacks on smart grids. Karimipour et al. [21] proposed an unsupervised anomaly detection method based on statistical correlation between measurements to distinguish between actual failures and interferences and intelligent network attacks. In the last two years, the power sector has also built frameworks or platforms based on AI algorithms, e.g., Tang et al. [22] built a three-dimensional framework of a large-scale power grid for digital simulation, and provided the knowledge model of power-grid simulation analysis and its application methods, i.e., the power-grid simulation knowledge-mining method and the AI model with the ability of diversified power-grid transfer learning. Sogabe et al. [23] embedded two deep-enhanced learning algorithms for discrete and continuous motion space design into a physical model, thus forming an algorithm library. Ferrag and Maglaras [24] devised a smart-grid energy framework based on deep learning and blockchain technology. Chi et al. [25] presented a heterogeneous multicore SoC hardware architecture containing a distributed processing unit (DPU). At present, the application mode of the fragmented effect of AI leads to poor technology portability, duplication of construction leads to waste in investment, and the lack of unified evaluation makes it difficult to guarantee the application effect. In addition, a large number of related research and development projects are being conducted that may be contradictory, resulting in data labeling, lack of standardization and generality, difference in interface standards, difficulty in measuring deep-learning training results, and other drawbacks. In view of these problems, it is necessary to build a standardized integrated and innovative application platform based on the common problems of intelligent power systems.

The AI platform for the power field from the viewpoint of functional positioning provides a model training environment and a production service environment; supports natural language processing and computer vision services; provides algorithms, models, and services for power business applications; and creates an ecological chain of "platform applications and services."

This paper provides an analysis of the deep learning framework, architecture design, key technology and application scenarios, and lays the foundation for the development and application of AI platform for the power field.

2 Deep Learning

Deep learning is a new branch of machine learning that has made breakthroughs in speech recognition, target detection, text processing, and more recently in AI [26]. Deep learning can be understood as a multi-hidden neural network, which realizes the approximation of any complex function by learning a deep nonlinear network structure and characterizing the distributed representation of data [27]. Deep learning focuses on the deep structure and characteristics of an automatic learning network, usually with 5, 10, or even hundreds of layers. A deeper hierarchy gives the network a stronger ability to automatically extract data features from a large number of samples.

Suppose we have system s, with n layers ($S1$, $S2$,$\cdots$ Sn), with input I and output O, it is represented as

$$I \Rightarrow S1 \Rightarrow \ldots \Rightarrow Sn \Rightarrow O \tag{1}$$

2.1 Deep Belief Network

Let us suppose a bipartite graph with no link between the nodes in each layer. In this graph, one layer is visible, i.e., the input data layer (*V*), and the other is hidden (*H*). If all nodes are random binary variable nodes (only 0 or 1 can be taken as values), and assuming that the total probability distribution *P* (*V*, *H*) satisfies the Boltzmann distribution, we call this system as the restricted Boltzmann machine (RBM).

$$p\left(h \mid v\right) = p\left(h1 \mid v\right) \cdots p\left(hn \mid v\right) \tag{2}$$

The deep belief network (DBN) is the first proposed deep-learning-network model consisting of several restricted RBMs.

$$p\left(v, h^1, h^2, \cdots h^l\right) = p\left(v \mid h^1\right) \cdots p(h^{l-2} \mid h^{l-1}) p(h^{l-1}, h^l) \tag{3}$$

2.2 SAE

The structure of an SAE is similar to that of a DBN and is stacked with multiple autoencoders (AEs). A basic AE can be thought of as a three-layer neural network with the same number of neurons as in the input layer.

An AE consists of an encoder and a decoder. Let us assume that the input of AE is

$$x = \left[x_1, x_2, \cdots, x_{dx}\right]^T \in R^{dx} \tag{4}$$

where *dx* is the dimension of the input. The encoder projects *x* from input layer to the hidden layer by mapping function *f*:

$$h = \left[h_1, h_2, \cdots, h_{dh}\right]^T \tag{5}$$

where *dh* is the dimension of the variable vector in the hidden layer, and function *f*(*x*) is expressed as

$$h = f(x) = s_f(W_x + b) \tag{6}$$

2.3 Convolutional Neural Networks

A convolutional neural network (CNN) is different from SAE and DBN in that it is a supervised, deep-learning-network model comprising an input layer, a convolution layer, a subsampling layer, a fully connected layer, and an output layer. The convolution and pooled layers are usually taken from several and are set alternatively.

A convolution is expressed as

$$S(t) = \int x(t-a) w(a) da \tag{7}$$

where asterisk represents convolution. A two-dimensional convolution is expressed as follows:

$$s(i, j) = (X * W)(i, j) = \sum_m \sum_n x(i-m, j-n) w(m, n) \tag{8}$$

In a CNN, a two-dimensional convolution is defined as

$$s(i, j) = (X * W)(i, j) = \sum_m \sum_n x(i+m, j+n) w(m, n) \tag{9}$$

The mapping of forward propagation is formulated as

$$x_{ij}^l = f(u_{ij}^l) = f\left(\sum_{p=1}^{s} \sum_{q=1}^{s} x_{i+p-1, j+q-1}^{l-1} \times k_{pq}^l + b^l\right) \tag{10}$$

and a convolution operation is given as

$$\frac{\partial L}{\partial K_{pq}^l} = \sum_i \sum_i (\delta_{ij}^l x_{i+p-1, j+q-1}^{l-1}) \tag{11}$$

2.4 Recurrent Neural Networks

Recurrent neural networks (RNN) are a type of neural network used to process sequential data; this is reflected in the calculation of the network's memory of the previous information and used in the calculation of the current output. The network is designed to have connections between neuron nodes of hidden layers, and the input of a hidden layer includes the input from the input layer and the output obtained from a previous task.

Output layer O and hidden layer S of an RNN are calculated as

$$O_{\mathrm{t}} = g(V_{S_t}) \tag{12}$$

$$S_{\mathrm{t}} = f(U_{X_t} + W_{S_{t-1}}) \tag{13}$$

2.5 Deep Reinforcement Learning

Deep reinforcement learning (DRL) is a combination of deep learning and enhanced learning. An end-to-end learning algorithm for perceived action, originally published as DeepMind, RL is a trial and error learning algorithm, which is more in line with human learning habits.

The Bellman equation is the core formula of RL:

$$V_\pi(s) = \sum \pi(a|s) E\left[R_{t+1} + \gamma V(S_{t+1}) \middle| S_t = s\right] \tag{14}$$

Further, the Bellman optimality is equated as

$$V_*(s) = E\left[R_{t+1} + \gamma \max_\pi V(S_{t+1}) \middle| S_t = s\right] \tag{15}$$

3 Architecture Design

The AI platform for the power field adopts a flexible hierarchical structure, and each layer is connected through a standard interface. The platform architecture design includes application architecture, technical architecture, data architecture, and deployment architecture.

3.1 Application Architecture

The application architecture of the AI platform for a power field is divided into five layers from bottom to top, as shown in Fig. 1: data lake storage, heterogeneous resource scheduling service, model development training warehouse, model reasoning, and platform management.

(1) Data lake storage integrates SQL, Non-SQL, and MPP databases to store multisource heterogeneous data,

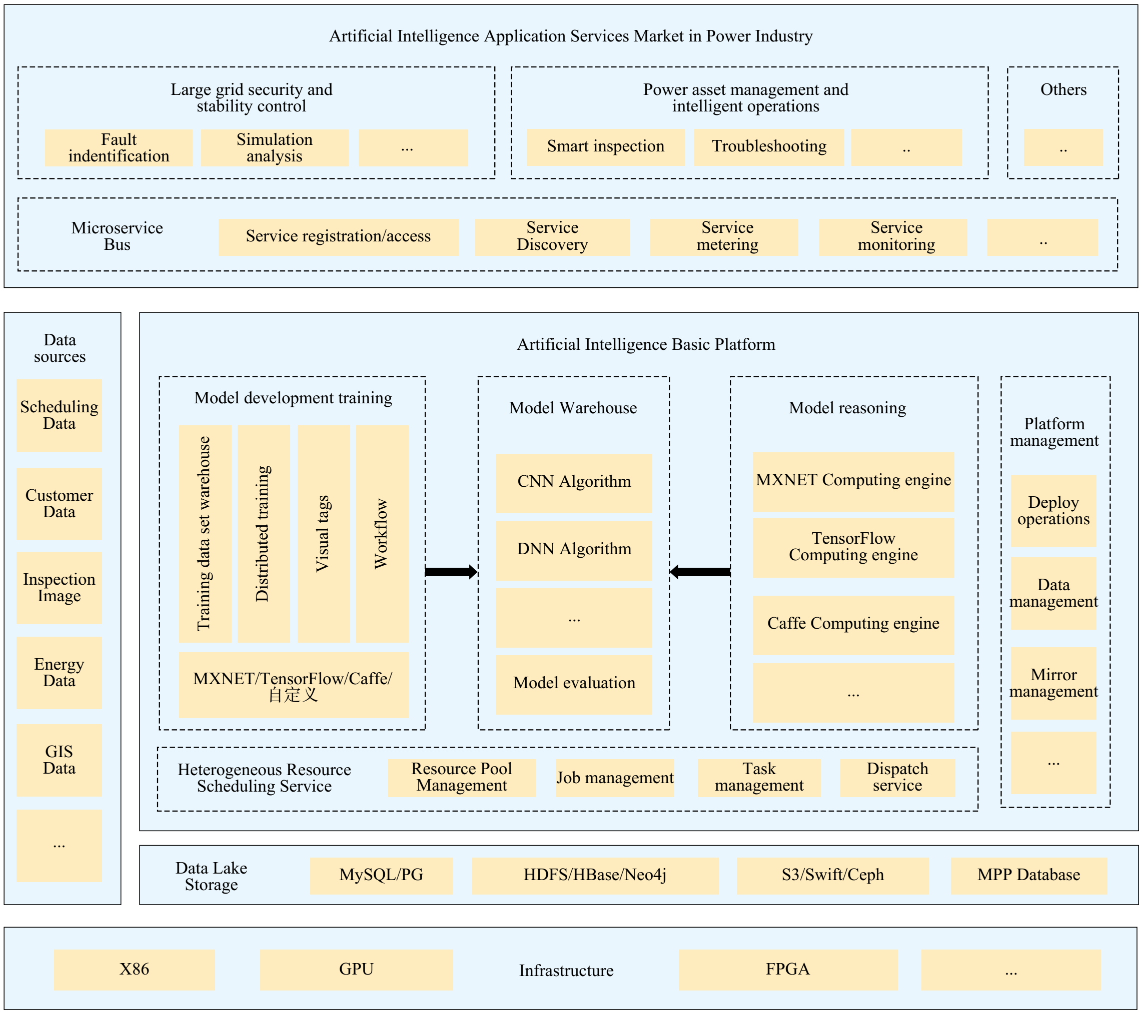

Fig. 1 Application architecture of electric AI platform

including power system scheduling data, customer service data, detection image, energy data, and GIS data.

(2) Heterogeneous resource scheduling services include resource pool management, job management, task management, and scheduling services.

(3) Model development training includes training dataset warehouse, distributed training, visual tagging, and workflow.

(4) Model warehouse includes a CNN algorithm, deep neural network (DNN) algorithm, and model evaluation.

(5) Model reasoning includes an MXNET computing engine, a TensorFlow computing engine, and a Caffe computing engine.

(6) The platform management includes deployment operations, data management, and mirror management.

3.2 Technical Architecture

The technical architecture of an AI platform for the power industries consists of five layers, i.e., data-acquisition, data-storage, data-processing, resource-scheduling, and service-performance layers, and two engines, namely, the deep-learning-algorithm and machine-learning-algorithm engines, as shown in Fig. 2.

(1) Data acquisition layer. The platform adopts multisource heterogeneous data, and data is collected using Flume, Kafka, Kettle, Sqoop, and more such tools.

(2) Data storage layer. To save the heterogeneous data collected, the platform uses various databases, e.g., the file storage uses NFS, the NoSQL database uses HBase/Neo4j, the relational database uses MySQL, the MPP database uses GBase, and the memory database uses Redis.

(3) Data processing layer. This layer processes data by using MR, Spark, Impala, data labeling, and other technologies.

(4) Resource scheduling layer. This layer uses a resource orchestration engine, a resource scheduling engine, Docker, Kubernetes, CUDA-Driver, and other technologies.

(5) Deep learning algorithm engine. This used the TensorFlow, Caffe, MXNet, PyTorch, etc.

(6) Machine learning algorithm engine. This engine uses Sklearn, Mllib, Python, R, and other technologies.

(7) Service performance layer. The layer uses intact and Vue, Spring boot, Istio, MyBatis, Traefik, etc.

3.3 Data Architecture

The sensor collects multisource data and places it in the data-lake unified storage pool for management. Through extraction, cleaning, transformation, and mining, the source data finally form various BI reports and unified knowledge map. Some data are labeled as data needed for model training, output models, and various AI services. The data architecture of an AI platform for the power field is shown in Fig. 3.

(1) Data from the storage layer are obtained through data acquisition, data integration, and other processes and placed into the data lake for unified storage.

(2) The data-labeling service in the compute layer calls the unified storage data in the data lake. GPU clusters are used for fixed and bulk labeling, and the results of data labels are stored into a data lake.

(3) Deep-learning services in the compute layer invoke data lakes to store data uniformly. GPU clusters are used to provide training and external services, and deep-learning results are stored in data lakes.

3.4 Deployment Architecture

The deployment architecture of the AI platform for power fields is divided into three layers, as shown in Fig. 4: internal and external network isolation, headquarters and directly affiliated companies, provincial and municipal companies.

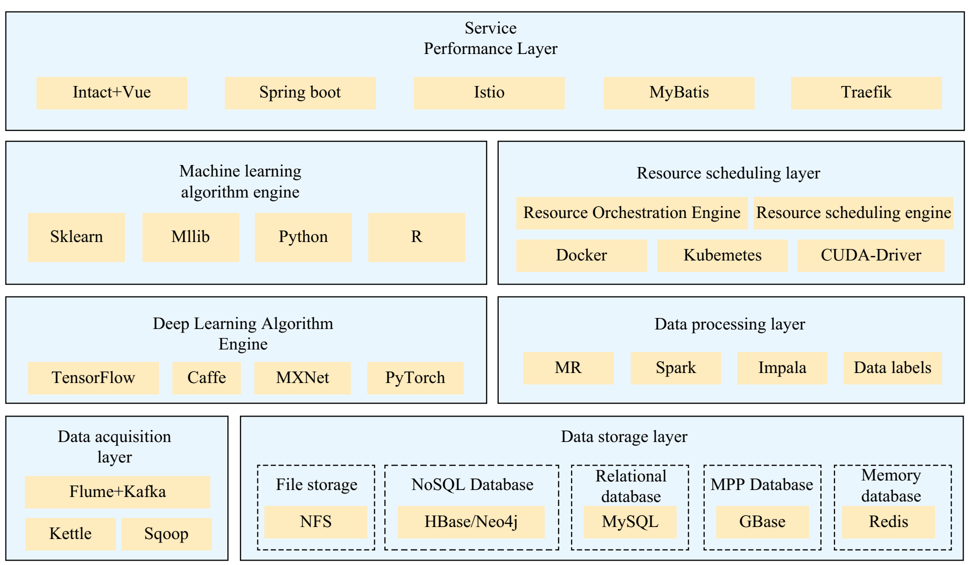

Fig. 2 Technical architecture of the electric AI platform

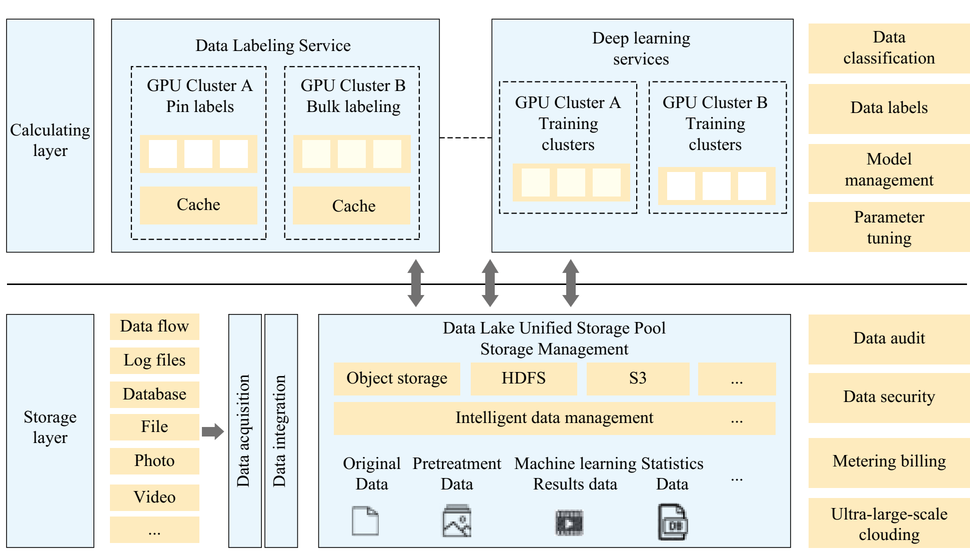

Fig. 3 Data architecture of the electric AI platform

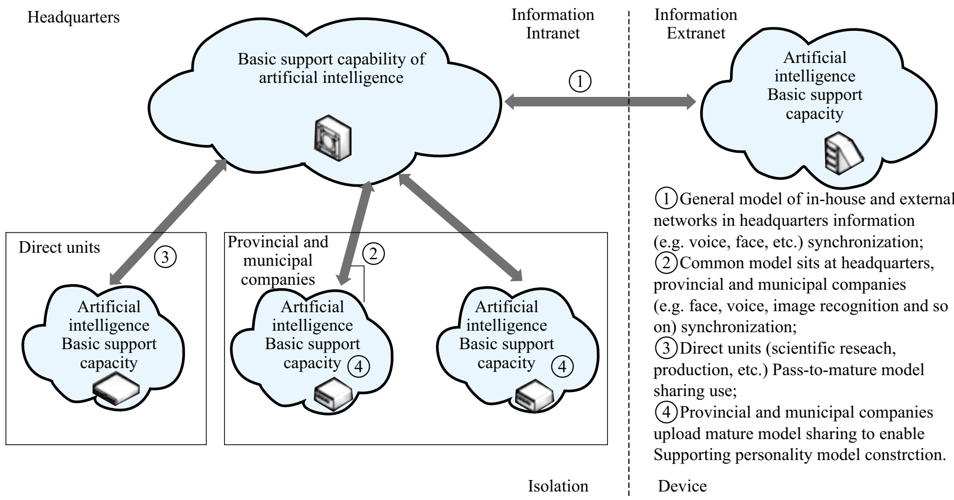

Fig. 4 Deployment architecture of the electric AI platform

The deployment architecture has the following capabilities.

(1) Basic support ability of AI in extranet involves deployment of information; providing voice recognition, face recognition, and other services; and synchronizing with the basic information of headquarters.

(2) Headquarters AI Foundation Support Capabilities are deployed in the information intranet through a unified operation and controlled from the headquarters of the National Network, complete with information extranet and direct units, including information interaction of provincial and municipal companies.

(3) The basic support capacity of AI in terms of direct units involves uploading a mature model to the headquarters and sharing its usage.

(4) The basic support capacity of AI for provincial and municipal companies involves the synchronization of the model shared with the headquarters, uploading of mature models to the headquarters, and support of personalized model building.

4 Key Technologies

4.1 Algorithm Model of Multihardware Compression and Conversion Technology

The algorithm model of the multihardware compression and conversion technology include the compression and acceleration of the original deep network model, and direct application of the million-level DNN model to the embedded device about mobile. The disadvantages of high storage and high power consumption promote the application of DNNs in applications with limited resources and real-time online processing.

(1) Parameter pruning

Parameter pruning removes redundant, less informative weights from existing trained deep network models, thereby reducing the parameters of the network model. Parameter pruning can speed up the calculation of the model, compress its storage space, and prevent it from overfitting. Parameter pruning can be divided into unstructured and structured pruning according to whether all the nodes must be deleted or filtered one at a time. Unstructured pruning considers each element for filtering and deletes the parameters, the number of elements of which is 0 in the filter. In contrast, structured pruning removes the entire filter of the convolutional layer without introducing additional data-type storage, thereby directly compressing the network and speeding up the calculation of the entire network.

(2) Parameter sharing

In parameter sharing, a mapping type is designed to allow sharing of the same data among multiple parameters; the methods of parameter sharing differ in several aspects. At present, the more widely used parameter-sharing expressions include parameter quantization, hash function, and structured linear mapping. Among them, the core concept of parameter quantization compression and accelerated deep network model is to replace the original 32-bit floating-point parameters with lower bits. Hash functions and structured linear mappings are used to achieve parameter sharing with the fully connected layer, greatly reducing the memory overhead of the model.

(3) Low-rank decomposition

The core concept of low-rank decomposition is to estimate and decompose the original convolution kernel in the depth model by using matrix or tensor decomposition technology. Convolution calculation shows the highest computational complexity in the entire CNN. The internal redundancy of the model is effectively reduced by decomposing the 4D convolution kernel tensor. In addition, the matrix parameters of the 2D fully connected layer can be processed using low-rank decomposition technology. However, owing to the different decomposition methods of the convolutional and fully connected layers, this project reviews and analyzes the application of low-rank-decomposition technology in DNNs according to these two layers.

(4) Design of compact convolutional filters

The design of compact convolution kernels mainly reduces the storage and computational complexity of models with the designing of special structured convolution kernels or compact convolution calculation units. Based on this concept, the original large filter size is directly decomposed into two 1 × 1 convolution filters, which greatly accelerate the network calculation and achieve high target-recognition performance.

(5) Knowledge distillation

Knowledge distillation mainly utilizes the knowledge of large networks and transfers this knowledge to the compact distillation model. The idea is to use the knowledge learned by a large model neural network (teacher network) a priori, and train a smaller neural network with better performance by transferring prior knowledge to a small-scale neural network. Therefore, the scale of the small network parameters after distillation would be much smaller than that of the original large network, and the purpose of compressing the network is be achieved. There is an intuitive way of transferring knowledge of a Big model (teacher network) by using the class probability vectors generated by the teacher network as soft targets for training the small model (student network).

4.2 Hybrid Heterogeneous Resource Distributed Scheduling Technology

The emergence of heterogeneous computing has solved the thermal and energy bottlenecks encountered in conventional methods to increase computing power by increasing the CPU clock frequency and number of cores. From an implementation perspective, heterogeneous computing formulates a series of software and hardware standards to allow different types of computing devices to share computing processes and results. At the same time, the calculation process is continuously optimized and accelerated to increase the computing efficiency.

(1) Distributed elastic scheduling technology based on container-based heterogeneous multisource algorithm

Container technology is a virtualization technology that has emerged in the last few years. It performs dynamic load balancing and elastic scaling based on Docker containers. When the application system faces high-concurrency and large-volume burst requests, the system monitors the CPU by using the application container. The usage of resources rapidly increases memory. The corresponding strategy involves setting up of the alarm threshold to trigger during the elastic scaling mechanism when the capacity is increased in steps; after the peak value is reached, the capacity is reduced by a fixed step size. When the application container is started or deleted, the corresponding key value of the elastic scaling module changes in Etcd. After Confd detects this change, it dynamically generates a Haproxy configuration file and reloads it. Next, it offloads front-end user requests to the back-end application container by providing specific services and implementing load balancing of the application system.

(2) Distributed-task-scheduling algorithm and system

Distributed task scheduling involves a framework that coordinates multiple nodes processing the same task to avoid data being repeatedly processed. The nodes of the cluster share the processing of large batches of tasks to improve the processing efficiency of batch tasks. The task-scheduling system completes task scheduling by quickly and efficiently distributing tasks, which will be executed with appropriate computing resources. In this process, the task-scheduling system ensures resource node load balancing, high-quality task-scheduling QoS, real-time updating of the task and resource statuses, monitoring of these statuses, task-scheduling error handling, and so on. The task-scheduling system plays an irreplaceable role in the entire task-scheduling process.

4.3 Automated Model Training and Evaluation Technology

Automated training technology (AutoML) is an end-to-end process of machine learning, and is proposed to automate all human tasks and reduce the threshold of machine learning. In typical machine-learning tasks, AutoML automatically learns by using the steps of modeling, optimization, and evaluation so that machine learning models can be applied without human intervention. The main problems of AutoML include the selection of the correct features, model family, and model parameters.

(1) Feature engineering

In machine-learning tasks, the quality of features largely determines the pros and cons of model performance. The automation problem of feature engineering involves the automatic construction of features based on data so that subsequent learning models show good performance. To achieve this goal, we primarily focused on the method of feature enhancement, that is, we performed some post-processing on these features through methods, such as dimensionality reduction, feature generation, and feature coding, to improve learning performance.

(2) Model selection

The model selection consists of two functions, namely selection of classifiers and setting up of the corresponding hyper parameters. In this AutoML setup, the task automatically selects the classifier and sets its hyper parameters to achieve a good learning performance. This requires a deep network architecture suitable for learning problems.

(3) Optimization algorithm selection

The optimization algorithm selection if focused on achieving good efficiency. However, as learning tools become more complex, optimization is not only a major consumer of computing budgets but can also affect learning performance. Therefore, the goal of algorithm selection is to automatically find an optimization algorithm to achieve a balance between efficiency and performance.

(4) Evaluation technology

After generating candidate configurations, the evaluator must measure their respective performances. In this case, the evaluator does not need to consider the configured search space. As such, the easiest way is to learn the model parameters and evaluate the performance.

4.4 Software and Hardware Acceleration Technology for the Algorithm Model

When the designer tries to achieve the best performance from the algorithm but the software method is no longer available, the designer could try to accelerate it through hardware/software repartition. Hardware acceleration refers to the usage of hardware modules instead of software algorithms to take full advantage of the fast characteristics inherent in hardware. From a software perspective, interfacing with a hardware acceleration module is the same as calling a function, the only difference being that this function exists in the hardware and is transparent to the calling function. Depending on the algorithm, the execution time can be up to 100 times faster. The use of hardware provides a much faster performance of various operations, such as calculation of complex mathematical functions, moving data from one place to another, and performing the same manipulation multiple times.

5 Application Scenarios

5.1 Scheduling Voice Visualization Scenarios

In the field of scheduling, a scheduling voice visualization scenario was constructed to provide services such as filling in of scheduling logs, managing scheduling services, and tracing fault problems. Scheduling voice visualization scenarios can free employees from heavy basic work, help improve work efficiency, and enhance safety and quality inspection of key tasks.

(1) Filling in the scheduling log

This involves automatically translating the dispatcher's calling content into text in real time, and copy pasting the relevant content into the dispatch log system after the call, eliminating the need for recording while the call is in progress.

(2) Scheduling business management

Here, the process of random inspection of recordings is converted into text records for inspection, and cases of noncompliance are confirmed through these recordings. In addition, we provide screening according to the department and seat to improve the pertinence of inspection.

(3) Backtracking of faults

Here, the dispatch recordings are converted into text; searched through keywords; and filtered according to time, unit, department, and seat. Finally, the corresponding recording is quickly located for investigation.

5.2 Drone-based Line Inspection

The fields of operation and inspection involve the inspection of transmission lines using drones and the provision of services, such as visually assisted autonomous line inspection, intelligent management of line inspection images, and intelligent recommendation of line inspection schemes. These help improve the reliability and efficiency of UAV-based line inspection.

(1) Visually assisted autonomous line inspection

Here, visual sensors are used to realize obstacle avoidance and autonomous patrol inspection based on wire tracking between poles. In addition, according to the target-recognition results of the airborne edge computing platform, the imaging parameters and drone attitude are automatically adjusted.

(2) Intelligent management of line-inspection images

The algorithm automatically eliminates duplicate and low-quality images, matches the line-of-sight images of the same scene at different times, and builds a multidimensional structured line-of-sight image library based on information, such as space time and tasks.

(3) Intelligent recommendation of line-inspection plan

According to historical line-inspection data and manual line-inspection experience, segmented quantitative evaluation of the line reliability level and automatic line inspection schemes are recommended. In addition, reliable line-inspection coverage is achieved at a smaller cost. Moreover, for emergency inspections in special situations, optimal inspection plan is recommended in a timely manner.

5.3 Metering and energy-analysis scenarios

The proposed framework sets up measurement and energy-analysis scenarios in the marketing field and provide services, such as electrical-energy-measurement equipment evaluation, energy analysis, and user portraits; this will help ensure the quality of the measurement equipment, promote the safe and reliable operation of user loads, and improve the user's power consumption experience.

(1) Evaluation of electric-energy-measurement equipment

To deeply explore the components of electrical-energy-measurement equipment and the impact of these components on the quality of the equipment, a component-comparison analysis robot was designed and developed. In the full performance detection and spot check of the electricity metering equipment, the component comparison and analysis can be used to compare the AI of the components, such as smart energy meters, power consumption information collection terminals, and measurement boxes, and avoid abnormal components. It can automatically identify the structural design and analyze the advantages and disadvantages of the comprehensive indexes to provide a reference for the design of electrical-energy-measurement equipment.

(2) Energy analysis

The AI algorithm-based user load identification and curve decomposition technology was adopted to construct a spatiotemporal database of typical pollution/emission reduction equipment waveforms for the entire industrial chain load, to achieve full information monitoring of typical industrial load equipment. Through the development of ubiquitous Internet-of-Things technology-based load intelligent identification micro-application functional modules for environmental management companies and selection of pilot provinces, in conjunction with the provincial environmental protection department in industrial parks, high-energy-consuming and highly polluting industrial enterprises could be selected for trial.

(3) User portrait

Based on massive data of electricity-consumption behavior, the model can identify the behavior characteristics of different customer groups, construct a predictive analysis model, analyze customer's energy-consumption characteristics, effectively predict customer electricity consumption risks, comprehensively analyze customer payment habits and preferences, and identify customer complaints hotspots. As such, accurate portraits of electricity customers could be identified to provide targeted value-added services to customers, optimize the electricity business environment, and improve customers' proactive service capabilities.

6 Conclusions

To break the data barrier, improve the portability of technology, and avoid investment waste caused by repeated construction, this paper proposes an AI platform framework for the field of power. The AI platform integrates three typical scenario applications, namely, scheduling voice visualization, unmanned patrol, and energy analysis.

The platform can meet the needs of current power system model training, natural language processing, and computer vision services. However, the platform faces technical difficulty in its integration with existing systems, such as EMS management, DTS, SCADA, and WAMS. As such, for the construction of this platform, we must determine how to unify the data interface, ensure integrity and confidentiality in the process of multisource heterogeneous data transmission, and manage the data stored in the platform.

At present, there are still some deficiencies in the platform, and these must be further strengthened in the follow-up research and development.

(1) The platform's data labeling function requires considerable amount of manual labeling. Therefore, the next step in this study is to correct the data labeling algorithm to reduce the workload to a certain extent.

(2) The spdin between platform models must be increased. Current models implemented by different deep learning engines cannot be run over other platforms, e.g., the Caffe model cannot be directly replaced with the TensorFlow model; therefore, the expansion of the model must be restricted to a certain extent.

(3) The model evaluation capabilities must be increased; at present, there is no effective evaluation mechanism after model training.

Acknowledgements

This work was supported by National Key Research and Development Project(2017YFE0112600) and Science and Technology Project of China Electric Power Research Institute(Research on the Key Technologies and Typical Application Scenarios of the Artificial Intelligence Basic Framework for Integrated Energy).

References

[1] Yang T, Zhao L, Wang C (2019) Review on Application of Artificial Intelligence in Power System and Integrated Energy System[J]. Automat Elect Power Sys 43(01): 2–14(in Chinese)

[2] Zhang B, Xu X, Xing H, et al. (2017) A Deep Learning Based

Framework for Power Demand Forecasting with Deep Belief Networks[C]. In: 18th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT). Taipei, Taiwan, pp 191-195

[3] Li P, He S, Han P, et al. (2018) Short-Term Load Forecasting of Smart Grid Based on Long-Short-Term Memory Recurrent Neural Networks in Condition of Real-Time Electricity Price[J]. Power Sys Technol 42(12):4045-4052 (in Chinese)

[4] Kumar S, Hussain L, Banarjee S, et al. (2018) Energy Load Forecasting using Deep Learning Approach-LSTM and GRU in Spark Cluster[C]. In: 2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT), Kolkata, India, pp 1-4

[5] Wang K, Zhong H, Yu N, et al. (2019) Nonintrusive Load Monitoring based on Sequence-to-sequence Model With Attention Mechanism[J]. Proceedings of the CSEE 39(01):75-83 (in Chinese)

[6] Wang S (2018) Research on Power Equipment Identification and Fault Diagnosis Based on Machine Vision [D]. North China Electric Power University

[7] He Z, Zhu F, Zhu Y (2018) Simulation of Fault Diagnosis of Transformer Fault in Power System[J]. Comp Simul 35(10):136-139 (in Chinese)

[8] Xu S, Qiu C, Zhang D He X, et al. (2019) A Deep Learning Approach for Fault Type Identification of Transmission Line[J]. Proceedings of the CSEE 39(01):65-74 (in Chinese)

[9] Shan X, Zhu B, Wang B, et al. (2018) Research on Deep Learning Based Dispatching Fault Disposal Robot Technology[C]. In: 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, pp 1-6

[10] Zhang P (2018) Research on Aerial Image Classification Based on Deep Learning Models[D]. North China Electric Power University

[11] Zhang Y, Qiu C, Yangfan, et al. (2019) Overview of Application of Deep Learning With Image Data and Spatio-temporal Data of Power Grid[J]. Power System Technology, 43(06):1865-1873 (in Chinese)

[12] Nguyen VN, Jenssen R, Roverso D (2019) Intelligent Monitoring and Inspection of Power Line Components Powered by UAVs and Deep Learning[J]. IEEE Power Energy Technol Sys J 6(1):11-21

[13] Li D, Wang X (2019) The Future Application of Transmission Line Automatic Monitoring and Deep Learning Technology Based on Vision[C]. In: 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, China, pp 131-137

[14] Cheng H, Han P, Dong Z, et al. (2018) Accurate Classification of UAV Aerial Images Contained Power Facilities[J]. Comp Simulat 35(08):424-428 (in Chinese)

[15] Hu W, Zheng L, Min Y, et al. (2017) Research on Power System Transient Stability Assessment Based on Deep Learning of Big Data Technique[J]. Power Sys Technol 41(10): 3140-3146 (in Chinese)

[16] Zhu Q, Dang J, Chen J, et al. (2018) A Method for Power System Transient Stability Assessment Based on Deep Belief Networks[J]. Proceedings of the CSEE 38(03):735-743 (in Chinese)

[17] Zhu Q, Chen J, Li H, et al. (2018) Transient Stability Assessment Based on Stacked Autoencoder[J]. Proceedings of the CSEE 38(10):2937-2946 (in Chinese)

[18] A Power System Transient Stability Assessment Method Based on Short-time Disturbed Trajectories[J]. Power System Technology, 2019, 43(05):1690-1697 (in Chinese)

[19] Wu L, Gao J, Venayagamoorthy GK, et al. (2018) On Artificial Intelligence Approaches for Contingency Analysis in Power System Security Assessment[C]. In: 2018 IEEE Power & Energy Society General Meeting (PESGM). Portland, OR, USA, pp 1-5

[20] Wang H, Ruan J, Wang G, et al. (2018) Deep Learning-Based Interval State Estimation of AC Smart Grids Against Sparse Cyber Attacks[J]. IEEE Trans Indus Inform 14(11):4766-4778

[21] Karimipour H, Dehghantanha A, Parizi R M, et al. (2019) A Deep and Scalable Unsupervised Machine Learning System for Cyber-Attack Detection in Large-Scale Smart Grids[J]. IEEE Access 7:80778-80788

[22] Tang Y, Huang Y, Wang H, et al. (2018) Framework for artificial intelligence analysis in large-scale power grids based on digital simulation[J]. CSEE J Power Energy Sys 4(4):459-468

[23] Sogabe T, Malla D B, Takayama S, et al. (2018) Smart Grid Optimization by Deep Reinforcement Learning over Discrete and Continuous Action Space[C]. In: 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC). Waikoloa Village, HI, USA, pp 3794-3796

[24] Ferrag MA, Maglaras L (2019) DeepCoin: A Novel Deep Learning and Blockchain-Based Energy Exchange Framework for Smart Grids[J]. IEEE Trans Eng Manag: 1-13

[25] Chi Y, Cui W, Qiao L, et al. (2019) Hardware Architecture Design of the Deep-learning-based Machine Vision Chip[C]. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Chengdu, China, pp 1110-1113

[26] Mohan N, Soman K P, Vinayakumar R (2017) Deep power: Deep learning architectures for power quality disturbances classification[C]. In: 2017 IEEE International Conference on Technological Advancements in Power and Energy (TAP Energy), Kollam, India, pp 1-6

[27] Zhu Y, Shi X, Wang L (2018) Recent Research Hotspot Introduction on the Application of Artificial Intelligence in Power System[J]. Power Gen Technol 39(03):204-212 (in Chinese)

[28] Zhou N, Liao J, Wang Q, et al. (2019) Analysis and Prospect of Deep Learning Application in Smart Grid[J]. Automat Elect Power Sys 43(04):180-197 (in Chinese)

[29] Ji R, Lin S, Chao F, et al. (2018) Deep Neural Network Compression and Acceleration: A Review[J]. J Comp Res Develop 55(09):1871-1888 (in Chinese)

[30] Zhu Y, Yin J (2018) Review of Research and Practice of Artificial Intelligence Application in Power Systems[J]. Power Generation Technology 39(02):106-111 (in Chinese). Biographies

Biographies

Peng Liu received Master's degree from North China Electric Power University, Hebei, 2014 and Bachelor's degree from Qingdao University of Science and Technology, Shandong, 2011. He is currently working at the China Electric Power Research Institute Co. Ltd., Beijing. His research interests include power big data platform, artificial intelligence technology, and Internet operation platform.

Wei Jiang received doctor's degree at Zhejiang University, Zhejiang, 2013. He is working in State Grid Corporation of China, Beijing. His research interests includes power big data technology, artificial intelligence platform and energy blockchain technology.

Xiaohui Wang received the Doctor's degree from North China Electric Power University, Beijing, 2012. He is currently working at the China Electric Power Research Institute Co. Ltd., Beijing. His research interests include power big data technology, artificial intelligence, active distributed network, energy internet.

HongMin Li received the Master's degree from North China Electric Power University, Hebei, 2013 and Bachelor's degree from North China Electric Power University, Hebei, 2010. She is working at the China Electric Power Research Institute Co. Ltd., Beijing. Her research interests include platform development and data visualization.

Hongjian Sun (IEEE, S'07-M'11-SM'15) received the Ph.D. degree in Electronic and Electrical Engineering from the University of Edinburgh, U.K., in 2011. He held post-doctoral positions with King's College London, U.K., and Princeton University, USA. Since 2013, he has been with the University of Durham, U.K., as a Reader in Smart Grid (with a Lecturer position in 2013-2017).

(Editor Zhou Zhou)