ISSN: 2664-9640

Architecture design of thermal processing cloud platform based on machine learning

Xianlong Peng, Baofeng Ji, Suchen He, Shizhong Wei, Guoshang Zhang, Hua Yu, Feng Mao

School of Information Engineering, Henan University of Science and Technology, Luoyang, China

Abstract

In view of the requirements of multi-source information fusion interactive process big data cloud platform for casting, forging, welding, and thermal basic manufacturing thermal processing processes, this paper proposes a machine learning-based thermal processing cloud service data sharing platform. Based on the construction of the big data cloud service platform architecture, the multi-source information fusion and interaction of data acquisition and feedback is realized, the process data of the whole process of thermal processing is effectively integrated, and a cloud service resource consumption model of thermal processing process data is established to realize Scheduling optimization of process data resources.

Keywords

Cloud platform, big data, machine learning, thermal processing.

1. Introduction

Industry is the lifeblood of a country's prosperity. With the introduction of advanced manufacturing strategies such as Made in China 2025 and Industry 4.0, new-generation information technologies such as cloud computing, big data, and the Internet of Things have been continuously integrated with industrial production, and related industrial applications have emerged as the times require. In the process of industrial production reform, how to improve the level of industrial production management, ensure the reliability of industrial processes, improve product quality, and visualize production process data hinder the development of industrial modernization, and the need to build a digital industry is imminent. The importance of technologies such as the Internet of Things, big data, and cloud computing in solving these problems is self-evident, but enterprises in the industrial industry are still exploring how to use these emerging technologies to better promote the development and advancement of industrial digitalization, and realize information. The premise of the application of technology to industrial development is the real-time and accurate acquisition of industrial process data. The integration of data acquisition with cloud services, industrial Internet and other technologies is the key to the digitalization of industrial processes.

Cloud computing provides three types of services: software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). Cloud manufacturing [1] is a new manufacturing model developed from the concept of "manufacturing as a service (MaaS)" derived from the Internet and Internet of Things (IoT), telecommunication networks, radio and television networks, wireless broadband and other networks. Based on the above three service models network of. Cloud manufacturing is a combination of existing manufacturing information technologies (such as computer-aided design, production, experimentation, simulation, management, integration) and emerging information technologies (such as cloud computing, Internet of Things, service computing, intelligent science, and efficient computing).

Virtualize and service various manufacturing resources and capabilities to form a service cloud pool for unified and centralized optimization management. The core of cloud manufacturing is to share manufacturing resources and capabilities. The key to cloud manufacturing is to build a public product platform for cloud manufacturing to achieve unified management of manufacturing resources and capabilities, and users can access on-demand through the cloud anytime, anywhere. So as to realize the efficient sharing and utilization of resources, and intelligently complete the whole life cycle of manufacturing [2]. With the continuous expansion of industrial production scale and the integration of production processes, the demand for industrial automation and digitalization has increased, the number and types of equipment in industrial processes have continued to increase, and the competition among manufacturing enterprises has become more and more fierce. The workshop production scheduling system improves its own competitiveness. Since the thermal processing big data cloud platform covers the requirements of multi-source cross-domain large-scale data acquisition and feedback, high concurrency sharing and secure access, multi-process parallel processing and computing, etc., this paper proposes a machine learning-based thermal processing cloud platform architectural design to realize resource scheduling optimization of processing technologies such as casting, forging, welding heat, etc.

The rest of this paper is organized as follows. The second section introduces machine learning and its application platform Hadoop for research, and the third section proposes a thermal processing cloud platform architecture design, and expounds the framework model of the machine learning-based intelligent manufacturing workshop intelligent scheduling platform. Finally, Section IV introduces the relevant outlook and concludes the paper.

2. Machine learning and its application platform Hadoop

In the context of the new model of intelligent manufacturing, the whole process data in the actual production process of the forging industry can be collected. Machine learning is the most core technology to realize data processing, data modeling, and make the production process intelligent. Machine learning is inseparable from big data, and its main purpose is to use algorithms to model and analyze the generated big data. Machine learning is mainly used in two aspects. The first type is business application data. Major Internet platforms can find out customer preferences according to a series of data such as customer purchase records and location information, and provide customized push messages to each customer, improve the return on sales investment. The second category is industrial data, which includes data on all aspects of production enterprises from raw materials to finished products. It is widely used in machine vision, deep learning, etc. in the industrial field, but it is less used in the direction of data processing.

2.1. Introduction to Machine Learning

Machine learning is a science involving artificial intelligence, probability theory, statistics and other fields. The main research goal of machine learning is how to find the model between the interrelated variables in many big data, and experience continuously improve the algorithm during learning, and continuously improve the algorithm by accumulating data. At present, my country's computer systems and industrial control systems have no learning ability, and at most can achieve some statistical analysis capabilities, which cannot meet the new needs of science and technology and production. In-depth research on machine learning will surely promote the further integration and development of artificial intelligence and the entire industrial field. As shown in Figure 1, machine learning can be divided into supervised learning, unsupervised learning and reinforcement learning according to the learning methods.

ISSN: 2664-9640

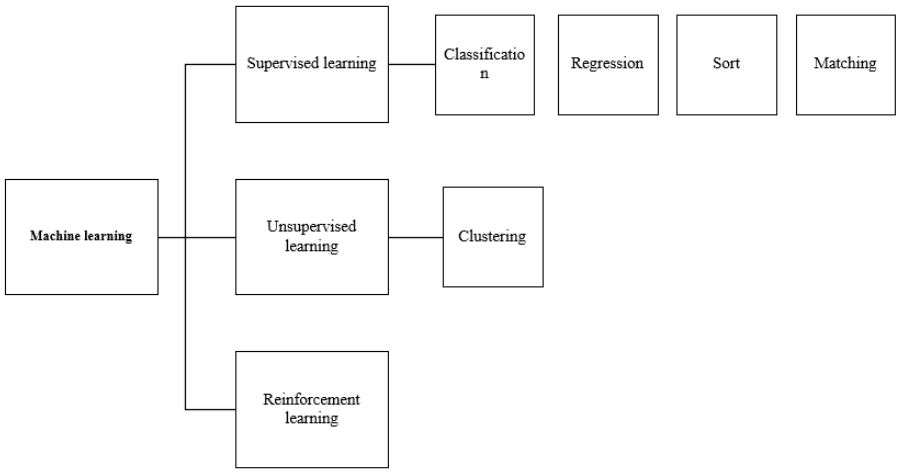

Fig.1 Classification of Machine Learning

(1) Supervised learning: The data has a certain clear label, and the algorithm is used to establish a mapping relationship between the input end and the output end, and predict the data model that may occur in the future. The meaning of the label is the category of the thing, and there are obvious differences between the labels. Currently commonly used supervised learning algorithms include logistic regression and statistical classification algorithms.

(2) Unsupervised learning: Unlike supervised learning, the data is not artificially given labels due to lack of experience and other reasons, and it is only divided according to the algorithm. Generally speaking, several types of representative data in big data are selected for testing. Common unsupervised learning algorithms include clustering.

(3) Reinforcement learning: Reinforcement learning does not belong to a machine learning algorithm of supervised learning and unsupervised learning. Reinforcement learning is composed of an agent (Agent) and the environment in which it is located. When the agent performs certain tasks, it will receive When the agent receives the reward, it will continue to update its knowledge base until the environment issues a termination command.

2.2. Hadoop framework

Hadoop is an open source framework for distributed systems developed by the Apache Foundation and used for big data computing and storage. Its biggest feature is a distributed system, which allows storage and processing in a distributed environment of computers using a simple programming model throughout the system. Big data, which also includes industrial big data. It is designed to scale from a single computing system to thousands of machines, each of which can provide local computing and storage.

The core technologies in the Hadoop framework are HDFS and Map Reduce. HDFS provides storage of massive data, and Map Reduce provides a series of methods such as machine learning to process big data. In the industrial domain, managers can independently develop Hadoop programs without knowing the details of the underlying equipment. Based on the advantages of high-speed computing in clusters, in another way, Hadoop is a software platform that can more easily develop and run industrial big data. The platform is implemented using the object-oriented programming language Java, which has good portability [3]. The following introduces the main functions of Hadoop, as shown in Figure 2.

ISSN: 2664-9640

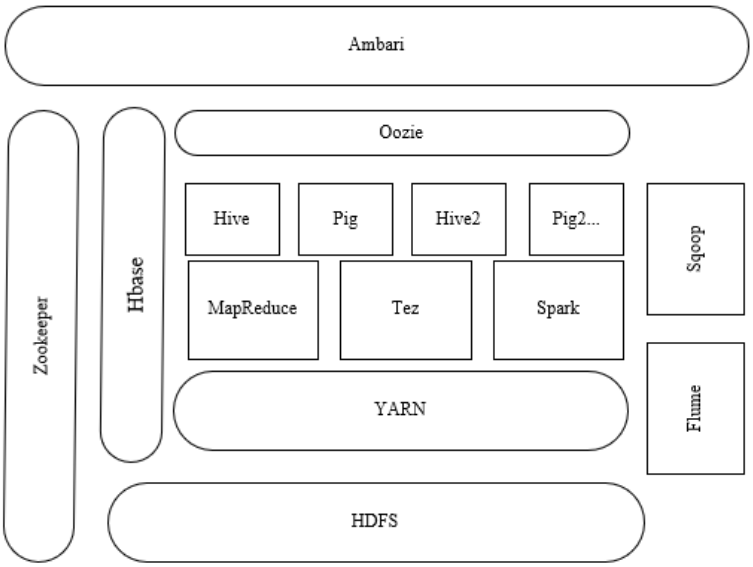

Fig.2 Key Features of Hadoop

HDFS: The definition of HDFS is a distributed file system. It has high fault tolerance and is suitable for various devices in the industry to provide high-throughput data access. Therefore, HDFS supports fast access to large files and read data. The operation speed is much faster than the writing operation speed. In view of the node aging caused by various types of failures in traditional industrial computers or servers, HDFS decomposes large files into 64MB data blocks, and uses the Linux file system .The decomposed data blocks are placed on each PC in the cluster, and redundancy judgment is set. In order to improve the redundancy of the system, the blocks belonging to the files are copied to each node in the cluster to ensure data security.

(2) Map Reduce: Map Reduce is a computing platform for parallel processing of big data. It can integrate thousands of clusters formed by common servers in the market at the same time, that is, Map Reduce connects the clusters. At the same time, Map Reduce is a parallel computing and The running framework means that he can divide computing tasks and computing data into various clusters to complete the parallel processing of tasks, which not only reduces the data processing time, but also greatly reduces the development time. At the same time, Map Reduce provides a method of parallel programming, that is, using Map to decompose the input Input into the intermediate Key/Value, and Reduce synthesizes the Key/Value into the final output Output.

(3) Hbase: Hbase is an open source, distributed database in Hadoop. It can quickly read and write a large amount of structured and unstructured data. In Hbase, data can be stored and retrieved arbitrarily to perform real-time data on the database. renew.

(4) Hive: Hive is also a data warehouse tool of Hadoop. The distributed database in Hive can provide a storage area for massive data, and it does not provide statistics and query functions in the SQL database language, and can export it into a data table. For example, data in HDFS can be exported [4-6].

ISSN: 2664-9640

3. Architecture design of thermal processing cloud platform

3.1. Overall Architecture Design

In order to adapt to the development of the situation and practice the concept of "Internet +", China Thermal Processing Corporation actively promotes technological innovation and application, researches and formulates new technology application standards and specifications in the thermal processing industry, and has launched Internet architecture cloud platform, big data platform and mobile technology. Manage the construction of platforms and other projects. The cloud platform realizes unified management and services of computing, storage, network, database, middleware and security; the big data platform completely covers data access, storage, processing and analysis, report presentation, data governance and other construction links. It manages multiple dimensions such as mobile devices, content, applications, and users, and realizes the visibility, management, and control of mobile office applications and data.

Cloud service data sharing is mainly composed of basic cloud platform, big data platform, cloud security, and cloud operation and maintenance. The cloud platform includes cloud infrastructure services (IaaS) and cloud platform services (PaaS). Cloud applications refer to related applications developed on the basis of cloud platforms according to user needs, and are an important part of the functions of cloud service data sharing platforms. Cloud security is the cornerstone of the entire cloud ecosystem. The cloud platform is the basic operating environment for various service components and information application systems. It consists of computing resources, storage resources, network resources, data collection platforms, and big data platforms, and has a complete operation and maintenance management mechanism and security guarantee system. The integrated cloud computing environment, its technical architecture is shown in Figure 3:

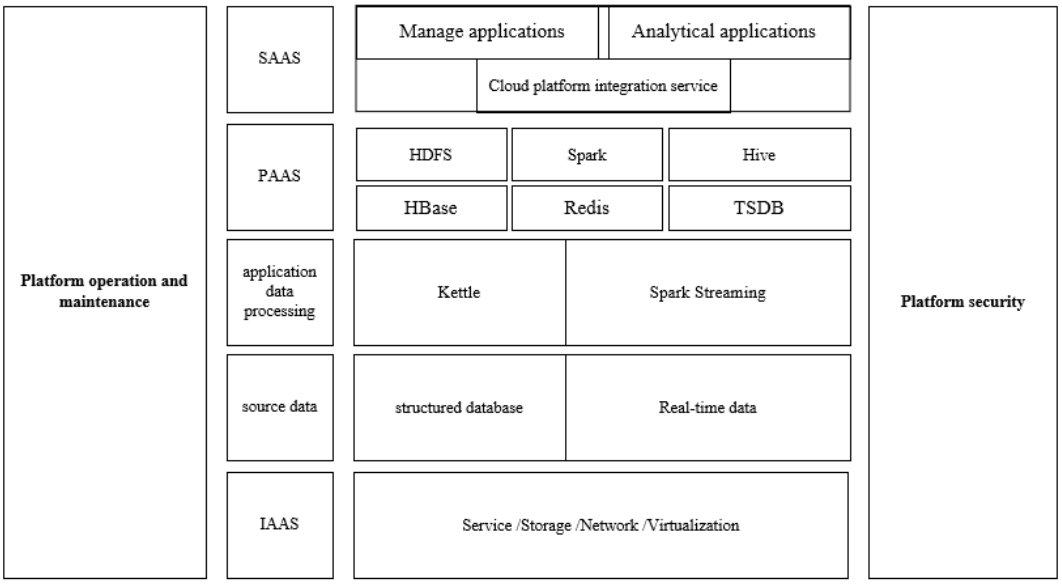

Fig.3 Technical Architecture Diagram

Platform Facility Layer

Platform IAAS layer mainly uses virtualization technology for physical resources and their services, so as to realize the pooled management of network, computing, storage and other resources, and can be flexibly allocated according to demand, for various business applications of CITIC Heavy Industry Provide complete cloud infrastructure services.

The platform component layer

The platform component layer is mainly data collection platform, big data analysis platform and mobile application platform. The data collection platform includes three parts: connection management, data management and analysis management. Among them, connection management is used for equipment registration management, remote access control, link strategy management, monitoring and alarm management; data management is used for data acquisition, data cleaning, and data storage. ; Analysis management for rules and alerts, predictive analysis. The big data analysis platform is mainly based on a set of mainstream high-efficiency computing platforms, including HDFS components, Spark components, Hbase components, Hive components, Redis components, and time series database components.

Unified Service

Based on each component of the platform, the unified service builds a unified cloud service data sharing platform management service system, including unified standards, that is, unified data collection standards; unified data, CITIC Heavy Industry data is collected into a unified data set; unified platform, CITIC Heavy industries share a unified platform for their respective business analysis; unified service, the platform provides unified information services for CITIC Heavy Industry.

3.2. Cloud Platform Design

Cloud platform, as the foundation and core of cloud service data sharing platform, is designed as follows:

1)The IaaS layer uses virtualization technology to logically form basic hardware resources such as computing, storage, and network to form a basic resource pool layer, and then re-encapsulates virtual machines, virtual storage, or virtual port groups provided by the resource pool with them. Combination, scheduling and use, form a virtual server or cloud storage system for organizational users, and provide resource services for each business unit of the thermal processing manufacturing system in this form; Full life cycle resource services.

2)The PaaS layer is based on the IaaS layer, based on the big data analysis platform, and uses key cloud computing technologies such as big data processing and mining analysis, intelligent applications, intelligent message push, social collaboration, and service-oriented architecture to process large-scale concurrency in real time. Load balancing and task distribution of tasks, so that all tasks are distributed and processed in real time, without accumulation, to achieve high reliability, no task will be lost in the process of processing, to ensure that all tasks can be processed, and to establish a cloud service data sharing ecosystem and Cloud service data sharing application ecosystem. Standardize, label and index various data resources, and provide data to various upper-level police applications in the form of services based on the information resource sharing platform. At the same time, it uses big data analysis and mining technology and information association technology to make deep use of data resources to form various thematic data resources.

3)The SaaS layer builds a unified cloud service data sharing platform management service system based on each component of the platform, which includes unified standards, that is, unified data collection standards; unified data, hot processing data is collected into a unified data set; Unified platform, the thermal processing process shares a unified platform for their respective business analysis; unified service, the platform provides unified information services for CITIC Heavy Industry and other information systems interaction.

4. Summary and outlook

In the context of intelligent manufacturing 2025, this paper takes the data in the thermal processing production process as the research object, and conducts in-depth research on the data sources, analysis methods and platform construction in the thermal processing production

process. The analysis method of big data in the production process is studied, and a platform for sharing data based on hot processing cloud service data is designed by using the machine learning algorithm platform Hadoop. It can greatly accelerate the informatization and intelligence of thermal processing, shorten the time and cost for optimizing the thermal processing process, and optimize the product quality. It is expected to reduce the process development cost by more than 50%. It is of great practical significance to shorten the process development cycle by more than 50%.

Future trends of cloud platform manufacturing services may cover the following studies: The depth of human-machine integration in the cloud service supply process, including the process of cloud manufacturing service platforms: The "push" or "pull" process also includes user needs, enabling user-centric, human-machine deep integration is based on the individual needs of R&D, manufacturing, and industrial organization, and promotes the formation of dynamic perception. The cloud-oriented personalized knowledge service method considers user behavior, integrates personalized demand characteristics into cloud manufacturing service research, and provides research ideas for deep human-machine integration.

Acknowledgements

This work was supported by National Key Research and Development Program (2020YFB2008400).

References

[1] Li BH, Lin Z, Lei R et al (2011) Further discussion on cloud manufacturing. Comput Integr Manuf Syst 17(3):449–457.

[2] Li B, Zhang L, Chai X (2010) Introduction to cloud manufacturing. ZTE. Communications 8(4):6–9.

[3] S.K.Lakshmanaprabu,K.Shankar,S.Sheeba Rani,Enas Abdulhay,N. Arunkumar,Gustavo Ramirez,J. Uthayakumar. An effect of big data technology with ant colony optimization based routing in vehicular ad hoc networks: Towards smart cities[J]. Journal of Cleaner Production,2019,217.

[4] Muhammad Babar,Fahim Arif,Mian Ahmad Jan,Zhiyuan Tan,Fazlullah Khan. Urban data management system: Towards Big Data analytics for Internet of Things based smart urban environment using customized Hadoop[J]. Future Generation Computer Systems,2019.

[5] Maria Malik,Katayoun Neshatpour,Setareh Rafatirad,Rajiv V. Joshi,Tinoosh Mohsenin,Hassan Ghasemzadeh,Houman Homayoun. Big vs little core for energy-efficient Hadoop computing[J]. Journal of Parallel and Distributed Computing,2019,129.

[6] Vanita Jain,Arun Kumar Dubey,Achin Jain,Mohit Malhotra,Shikha Rastogi. Crime pattern recognition in Chicago city using hadoop multinode cluster[J]. Journal of Information and Optimization Sciences,2019,40(2).