

Intelligent Analysis and Application of NLP and OCR Technologies in Power Grid Project Evaluation

Jiangqian Huang^{1,*}, Yihui Wang², Lin Jin², Yuhang Wang²

¹Economic Research Institute of State Grid Zhejiang Electric Power Company, Hangzhou, China

²Zhejiang Huayun Information Technology Co., Ltd., Hangzhou, China

*Corresponding author: 49767462@qq.com

Abstract: With the construction of socialist economy and the continuous improvement of national economic level, stricter requirements have also been placed on power grid engineering. The intelligent evaluation technology in China's power grid engineering has made significant breakthroughs so far. This article uses NLP and OCR technology for intelligent analysis in the evaluation of power grid projects, and analyzes the application of this technology in it. Combining NLP technology for adaptive learning and error compensation in the process of extracting key indicator features for distribution network projects, extracting refined indicator parameters for human, financial, and material aspects of distribution network projects to achieve intelligent extraction of key indicators for distribution network projects. OCR technology is the most critical technology in the process of digitizing paper text archives, which changes the traditional concept of inputting paper media data. This article effectively improves the information feedback speed of power enterprises through NLP and OCR technology, fully ensuring the stability, environmental protection, efficiency, and economy of electricity supply. The technology presented in this article has significantly improved the operational efficiency of the power grid and improved the operational efficiency of power enterprises.

Keywords: Power grid project review, NLP, OCR, intelligent analysis, application

1. Introduction

Technological innovation and technological progress are the inevitable products of the development of the times. Nowadays, electric power has become an indispensable key energy source in people's lives. The development of electric power technology and power grid technology has greatly affected social construction and people's livelihood needs. China is facing an important problem of shortage of energy resources, and the rising population under the background of globalization has aggravated the supply burden of energy resources on earth [1]. Although China has a vast territory and abundant resources, it has a large population and a small amount of resources per capita. As one of the necessary energy sources for people's production and life today, the importance of electric energy cannot be ignored [2]. With the increasing demand for electricity, the problem is the shortage of electricity and other resources. How to effectively improve the shortage of power resources and improve the utilization rate of power resources is the direction that needs to be focused in the research of power engineering technology [3-4]. In this paper, NLP and OCR technology are used for intelligent analysis in power grid project evaluation, and the application of this technology is analyzed. In order to lay a model foundation for extracting key indicators of distribution network projects, a distributed big data mining model for distribution network projects is designed based on the engineering data center and with intelligent perception, and big data mining and adaptive scheduling are carried out by changing the contents of data packets, changing the ID of data packets and forwarding data packets [5]. Combined with NLP technology, adaptive learning and error compensation are carried out in the process of key index feature extraction of distribution network project, and the refined index parameters of human, financial and material of distribution network project are extracted to realize intelligent extraction of key index of distribution network project [6]. OCR technology is the most critical technology in the process of realizing the electronization of paper files, which changes the traditional concept of data input in paper media. Compared with the traditional manual input method, OCR technology can greatly improve the efficiency of data storage, retrieval and processing [7]. NLP and OCR technology can effectively improve the information feedback speed and information transmission speed of power enterprises, which is more conducive to improving the quality and function of power services and fully ensuring the stability, environmental protection, high efficiency and economy of power supply [8].

2. Intelligent extraction and optimization of key indicators of power grid projects

2.1 NLP technology

Intelligent review technology is also a branch of smart grid technology. Smart grid refers to a new type of power grid with high technological content. It is a system control for analyzing power grid projects by adding technology and automation to the existing physical power grid. This chapter utilizes NLP technology to intelligently extract and optimize key indicators of power grid projects. Based on the data structure characteristics of key indicator big data, this paper proposes an intelligent extraction model for key indicators of power grid projects based on NLP[9]. The smart grid has three characteristics: economy, intelligence, and durability. Firstly, it is about conservation. As our country has a large population, the per capita possession of various resources is relatively low. Therefore, in this situation, the efficient and practical utilization of per capita resources is the focus of China's resource economy development. With the support of NLP technology, the minimum support is defined as:

$$\min EspSN = |D| \times \eta \quad (1)$$

Under NLP learning and training, the training time of NLP is defined as $Timer(Task[i], Source[j])$, which means that the extraction task of the key indicator i of the power grid project is defined as $Task[i]$, and the time required to allocate the extraction task to the resource $Source[j]$. The completion time of any key indicator feature extraction scheme of the power grid project is expressed as:

$$Timer(x) = \max \left(\sum_{i=1}^k (Task[i], Source[j]) \right) \quad (2)$$

Adaptive learning and error compensation during the feature extraction process of key indicators in power grid projects, with the learning weight of $w^{(k)}$ for NPL and the output feature quantity meeting the following requirements:

$$D(s) = 1 + C_2(s)P(s) = 1 + \frac{\Delta Ke}{(\lambda_2 + L_m)} \quad (3)$$

Adopting NLP method and applying big data analysis technology in the design process to enrich the general design scheme of power grid projects. Develop specific power supply methods based on the type of facilities in the power grid. Devices that can be accurately configured should be able to effectively avoid energy waste and component loss caused by power mismatch [10]. Realize intelligent analysis of cost and material information.

2.2 Intelligent extraction and output of key indicators for power grid projects

Intelligent evaluation is also indispensable in the transmission of power grids. The smart grid must have sufficient stability in the process of power generation and transmission, which requires higher technical operations. The application of power engineering technology in power sources can help the power sector provide targeted power sources, such as using charging technology for battery charging, and using high-frequency switching power sources for large electronic equipment charging[11]. There are almost strict requirements for equipment requirements and data recording, which not only require the integration of a large number of advanced technology and more high-end facility spare parts, but also the most rigorous and meticulous assessment of power grid staff. In this regard, this article applies 3D visualization interaction technology to develop standardized construction tool software for three project departments that can be flexibly configured, improving the object-oriented nature of key indicator extraction. Key indicators of distribution network projects: In distributed sensor storage media, the output load is:

$$g(t) = \frac{1}{\sqrt{a_0}} f \quad (4)$$

Assuming $t_0 = f_0 T / B$, the center frequency of information collection for key indicator feature extraction in distribution network projects is f_0 . By initializing the data chain from the cluster center to the points within the cluster, the Doppler tolerance for key indicator extraction in distribution network projects can be obtained as follows:

$$\xi(w) = \min\{\Pr(T \leq \zeta)\} \quad (5)$$

The link random allocation method is used to control the balanced output sensing sequence of big data, the key index of distribution network project. According to the above algorithm design, the refined index parameters of human, financial and material of distribution network project are extracted. The power grid also contains rich intelligent technology and scientific technology, which can effectively reduce the use of human resources and materials in energy development and further strengthen the efficiency of solving various faults. Combined with big data mining, Internet of Things, intelligent perception, artificial intelligence and other technologies, intelligent extraction of key indicators of distribution network projects is realized.

3. Application of OCR technology in power grid project evaluation

3.1 System structure

The terminal of power grid operation is the link of electricity consumption. Now, users no longer need to rely on the data displayed on the electricity meter to determine the actual electricity consumption, which is a method with too large precision difference and relatively single function, but use the more precise, precise and diversified services of smart grid to make the calculated electricity stable and prosperous. This function of smart grid enables users to query power consumption information at any time regardless of location and time. After receiving the certificate information submitted by the user, the counter assistant shoots the certificate information through the overhead camera, and the system uses OCR technology to automatically identify the information items in the application form and automatically fill in the application form. The system design adopts a modular structure, and each functional module will not only be used for single business, but also be applied to other business applications. At the same time, the OCR identification module of each certificate is customized, developed and integrated to realize the intelligent collection of information and the user's exemption from filling in the form. The OCR identification module of residents' renaming and transfer mainly includes the ID card collection and identification module, the household registration book collection and identification module, the real estate license collection and identification module, and the real estate title certificate collection and identification module. Each module also provides information interfaces to provide intelligent information collection and identification support for related business systems. The structure of the business application system is shown in Figure 1.

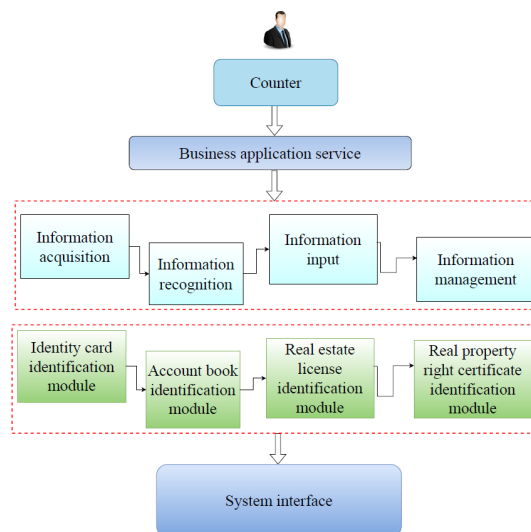


Figure 1: Business Application System Structure

The combination of power engineering technology and smart grids has also led to the development of automation in the management function of power dispatch. So, the intelligent evaluation system plays a promoting and supervising role in this change, and is also a prerequisite for ensuring the stable development of the power industry, making it more convenient for people's electricity consumption and life. After completing OCR recognition, the system obtains the text in the uploaded document image and the corresponding relationship between the text and the document layout item. At this time, the system can automatically complete and publish the form for users, generate text content according to the rules corresponding to different documents, fill it in the front-end corresponding publishing form, and automatically publish it. The system will code the ID images to prevent their information from being leaked and then released.

3.2 Operation flow

The user shall provide proof materials to the salesperson, including ID card, household registration book, property ownership certificate or real estate certificate, and inform the relevant information of the household number and original household name. The counter salesperson uses resident renaming and transfer business applications to collect and identify documents through a high-speed camera. The OCR recognition application functions for various documents have been integrated into the system, and the document information is automatically filled in the digital application form after recognition. After customizing the OCR application for document information collection and recognition, the OCR collection information sharing interface can provide standard interfaces to share various types of document collection information, providing information or functional support for other business applications that require the same document OCR recognition. The system will check the received image files, check the validity of the image file format and the size of the images, then compress the images, code key information locations, and finally rename and upload the processed images to the server, returning the processing results. The logical design diagram of the OCR recognition module is shown in Figure 2.

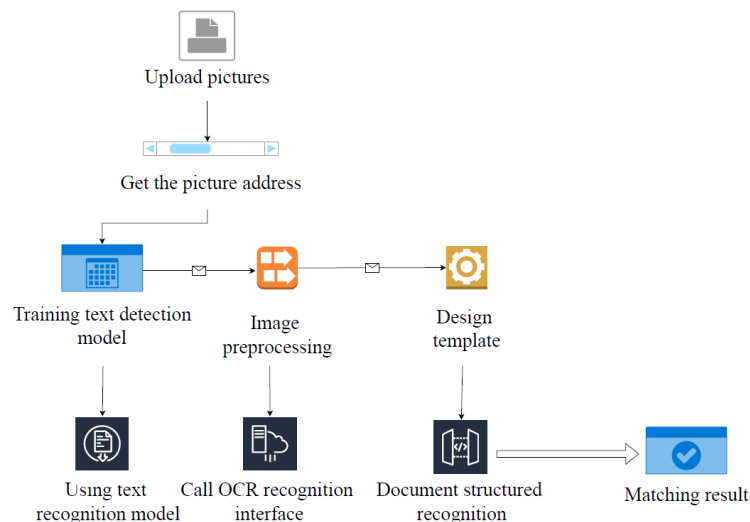


Figure 2: OCR recognition logic design diagram

When uploading and publishing information and uploading images, the image file should first be uploaded to the server graph bed, and the format, size, and other attributes of the image should be checked. The image address that meets the requirements will be passed to the preprocessing code of the OCR text recognition module for processing. After confirmation by the salesperson, the application form is submitted, and the application information is pushed to the marketing system. The salesperson submits the application after confirmation in the business processing module of the marketing system, and the business processing is completed.

4. Conclusions

Nowadays, the continuous progress of economic construction is also promoting the continuous progress of power engineering. Intelligent assessment technology is more widely used in smart grid, and smart grid is in an extremely important position in the whole power system. In this paper, NLP and OCR

technology are used for intelligent analysis in power grid project evaluation. NLP method is adopted, and big data analysis technology is applied in the design process to enrich the general design scheme of power grid projects. According to the type of power grid facilities, specific power supply methods are formulated, DC and AC are set, and instruments that can be configured accurately should be able to effectively avoid the waste of energy and the loss of parts caused by power mismatch. Using OCR technology to automatically identify the information items in the application form and automatically fill in the application form, the system design adopts a modular structure, and each functional module will not only be used for single business, but also be applied to other business applications. The business model, system structure, business process and OCR intelligent collection mode standard realized by OCR technology can change the original traditional business application processing mode of power supply business window, eliminate the link of filling in the form during the processing, and reduce the information input of front-line staff. NLP and OCR technology can be used in the evaluation of power grid projects to control the power consuming objects in the power grid and quickly process and analyze the data in order to get the most accurate feedback control signal.

Acknowledgements

We acknowledge funding from the 2023 Comprehensive Plan Special Project of Economic Research Institute of State Grid Zhejiang Electric Power Company (Grant no. JY02202325).

References

- [1] Wang Q, Li H. *Application of Power System Intelligent Control in Data Analysis of Power Grid Dispatching Control System*[J]. *2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI)*, 2022, 28(11):1037-1040.
- [2] Zheng C, Zhang C, Liu Q, et al. *Analysis on application scheme of electronic fence technology in power grid infrastructure project based on Beidou Navigation Positioning Technology*[J]. *Journal of Physics: Conference Series*, 2021, 1871(1):012065-012072.
- [3] Liu Q, Li T, Zang Q, et al. *Research on key technology of 3D digitization for secondary system of power grid based on model application analysis*[J]. *Journal of Physics: Conference Series*, 2021, 17(2):15-26.
- [4] Liang F, Feng Q I, Lu-Jie Q I, et al. *Design of Business Collaboration Oriented Power Grid Planning and Intelligent Decision Platform*[J]. *Electric Power Information and Communication Technology*, 2018, 17(7):36-44.
- [5] Oh Y K, Yi J. *Asymmetric effect of feature level sentiment on product rating: an application of bigram natural language processing (NLP) analysis*[J]. *Internet Research*, 2021, 36(21):20-29.
- [6] Ou C. *Application analysis of intelligent construction technology in power grid construction project*[J]. *Peak Data Science*, 2017.
- [7] Shao G, Wang B, Liu Y, et al. *Key Technologies of Multi-Source Data Integration System for Large Power Grid Operation*[J]. *E3S Web of Conferences*, 2021, 26(10):10-14.
- [8] Ling-Fan L I, Bei-Ning S, Jie Du. *Application of Remote Review Platform in the Power Grid Informatization Project Review*[J]. *Techniques of Automation and Applications*, 2019, 31(11):22-30.
- [9] Hu Z. *Power Grid Planning and Dynamic Intelligent Management System Construction Based on Operation Simulation* [J]. *Journal of Physics: Conference Series*, 2021, 2143(1):012015-012021.
- [10] Xuan Yu, Yang Su, Weiyang Z. *Research and application of automatic operation and maintenance in safety management and control platform in power grid enterprises based on big data analysis*[J]. *Power Systems and Big Data*, 2018, 22(12):36-42.
- [11] Zhao M, Wang Z, Xue Y. *Introduction an overview on application analysis of power electronic technology in smart grid* [J]. *Power Systems and Big Data*, 2018, 12(5):10-22.