Development of College Entrance Examination Voluntary Application Assistant System Based on Big Data Technology

Haidong Yang

Philippine Christian University Center for International, Education, Manila, 0900, Philippines;
Nanchang Institute of Technology, Nanchang, Jiangxi, 330044, China
[*]Corresponding author's e-mail: xdfc8562@126.com

Abstract

The reform of college entrance examination is a major measure of the reform of talent selection mechanism. It not only pays attention to the comprehensive quality of students and respects the right of students to choose independently, but also brings difficulties for examinees and families to fill in the application for college entrance examination. It has become the focus of social attention how to make full use of big data in previous years to guide voluntary application and reduce the difficulty of voluntary application through intelligence. In this context, it is particularly important to design and manage the voluntary filling system based on the new college entrance examination reform. In order to enable the examinees to make a comprehensive reference to the scores and rank of each university and major over the years, to better position their own scores, and then rationally select majors and college aspirations, and increase the chances of being admitted to the ideal university and major, it is necessary to refer to the intelligent voluntary application decision-making system based on big data. This paper mainly describes the development and research of the college entrance examination voluntary filling assistant system based on big data technology, aiming at strengthening students' attention to the college entrance examination voluntary filling and helping students make comparison and decision-making.

Keywords: *Big data technology, voluntary filling auxiliary system*

1 INTRODUCTION

With the development of the Internet era, the number of websites and mobile applications has gradually increased. With the reform of the new college entrance examination system, the issue of education has attracted much attention. Promoting educational equity and helping students to carry out educational career planning are related to the training and development of national teenagers, which has also received a positive response from the public [2]. With the continuous changes of the college entrance examination voluntary filling system, the filling method has changed from "filling before the examination" to "estimating the score filling" to "knowing the score filling". This has certain advantages for students [3].

2 RELATED TECHNOLOGY INTRODUCTION

The first is the key technology of big data processing. Key technologies of big data processing mainly include big data collection, big data pre-processing, big data storage and management, big data analysis and big data visualization. Big data collection mainly refers to the acquisition of various types of massive data from the network, which is the fundamental of big data data processing. Big data collection can be achieved by data crawling technology. Big data preprocessing refers to data discrimination, extraction, cleaning and other operations. Through big data preprocessing, complex data can be converted into data structures that are easy to process [4]. For some data that we do not care about, we can remove them, so as to achieve rapid analysis and processing [5]. Big data storage and management requires sufficient space to store data. In general, the HDFS of Hadoop can be used to store big data. Big data

B. Fox et al. (Eds.): IC-ICAIE 2022, AHCS 9, pp. 1084-1089, 2023.
https://doi.org/10.2991/978-94-6463-040-4_163

analysis refers to the use of big data processing technology to break through user interest analysis and network behavior analysis. It can use Hive, HBase and other technologies in Hadoop ecosystem and combine some big data processing algorithms for data analysis. Big data visualization refers to the visualization of data analysis results to users, so that users can see the analysis results intuitively, such as graphical display, web display, etc [6].

The second is Hadoop platform technology. Hadoop is a reliable, scalable distributed system infrastructure for distributed computing and distributed open source software developed by the Apache Foundation. The core module of Hadoop is divided into two parts: HDFS and MapReduce. The Hadoop Distributed file system (HDFS) is an extensible, fault-tolerant, and high-performance Hadoop distributed file system. It is used for storing and backing up distributed files. The HDFS consists of one NameNode and multiple Datanodes. Namenodes are used to store metadata and process requests sent by clients [7]. Datanodes are where data is stored. NameNode is responsible for recording the file block information. The storage space of HDFS is determined by the storage space of Datanodes. MapReduce is a distributed computing framework of Hadoop, which consists of mapping and specification processes [8]. The core idea of MapReduce is as follows: When a MapReduce job is started, the Map end reads data from the HDFS, maps the data into key/value pairs, and sends the data to the Reduce end. The Reduce end receives key/value pairs sent from the Map end, groups the data based on different keys, processes the data with the same keys, generates new key/value pairs, and outputs the result to the HDFS [13]. A complete MapReduce process consists of four stages: reading input data, and each data block in HDFS generates a Map; Map phase: Data is read in the form of key-value pairs, processed as required, mapped into new key-value pairs, and sent to the Reduce end. In the Shufflfffffle/Sort phase, data with the same keys output in the same Map is integrated in a step to reduce the amount of transmitted data, and the data is sorted by key after integration [9]. In the Reduce phase, the number of Reduce tasks is determined based on the data partitions set in the Map phase. Data in a partition is processed by a Reduce task. Data with the same keys is specified in each Reduce task processing and output to the HDFS in a new key/value pair [10].

Table 1 Related technology introduction

Related technology introduction	
The key technology of big data processing	Hadoop platform technology
Key technologies of big data processing mainly include big data collection, big data pre-processing, big data storage and management, big data analysis and big data visualization.	The core module of Hadoop is divided into two parts: HDFS and MapReduce.
Big data collection mainly refers to the acquisition of various types of massive data from the network, which is the fundamental of big data data processing.	A complete MapReduce process consists of four stages: reading input data, and each data block in HDFS generates a Map; Map phase: Data is read in the form of key-value pairs, processed as required, mapped into new key-value pairs, and sent to the Reduce end.

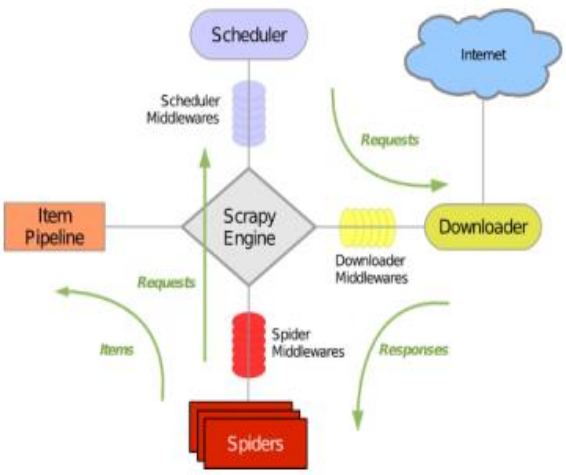

Figure 1: Scrapy frame components diagram

3 SUPPORT VECTOR MACHINE TO ESTABLISH THE COLLEGE ENTRANCE EXAMINATION VOLUNTARY REGISTRATION MODEL

Support vector machine (SVM) is a classification algorithm among machine algorithms. The classification goal is to find an optimal classification hyperplane, distinguish all the data of college entrance examination, and maximize the distance between all samples and the classification hyperplane [11]. Let the voluntary filling data of college entrance examination be: {x1, y1}, x1∈ Rn, I = 1,2,3... N, X represents the input of the college entrance examination, Y represents the risk level of the college entrance examination, so it can be described by the following formula:

$$Y = W^{T}\varphi\ (x)\ + b \tag{1}$$

Where W represents the offset value of the optimal classification plane, so it can be obtained:

$$\min J\ (\omega,\gamma) = \frac{1}{2}W + C\epsilon\gamma i \tag{2}$$

In order to further obtain the solving process, it is necessary to speed up the calculation. The specific formula is as follows:

$$\max W\ (\alpha) = \epsilon ai\text{-}\frac{1}{2}W\gamma \tag{3}$$

$$\text{Tab}\quad P(a)\quad P\ (b) \tag{4}$$

The chart of risk assessment model established by support vector machine is as follows:

Table 2: Schematic diagram of the chart

	x1	x2	x3	x4
1	0.055	0.630	0.942	0.542
2	0.211	0.312	0.432	0.894
3	0.537	0.438	0.088	0.670
4	0.131	0.581	0.853	0.548
⋮	⋮	⋮	⋮	⋮
500	0.475	0.542	0.324	0.590
501	0.360	0.348	0.942	0.846
502	0.871	0.482	0.058	0.518
503	0.956	0.651	0.632	0.040
504	0.701	0.921	0.531	0.994
⋮	⋮	⋮	⋮	⋮
1 000	0.888	0.654	0.420	0.300

The steps of establishing the risk assessment model of college entrance examination voluntary registration by support vector machine are as follows: firstly, the risk assessment indicators of college entrance examination voluntary registration should be collected and analyzed, and the risk assessment indicator system of college entrance examination voluntary registration should be established through experts. Ahp is used to determine the index hierarchy of the voluntary examination for college entrance examination, and grey correlation analysis is used to screen the risk assessment indicators of the voluntary examination for college entrance examination, and select the indicators that have an important impact on the risk assessment results of the voluntary examination for college entrance examination. Support vector machine is used to study the training samples of risk assessment for college entrance examination voluntary registration, determine the corresponding parameters, and establish the risk assessment model for college entrance examination voluntary registration [12].

4 ANALYSIS AND IMPROVEMENT METHODS OF APP PROSPECT OF COLLEGE ENTRANCE EXAMINATION APPLICATION

The first is market research, which requires research on the target market and similar markets. According to the market orientation studied in the questionnaire survey, market research was conducted on the characteristics of the consumer market of senior high school students and the consumption situation of parents, and the initial market report was formed [1]. The report focuses on the future consumption orientation of high school students and their parents as well as their willingness to accept "platform information services" and their requirements for subjective and objective information. Secondly, it is necessary to conduct an investigation on the customers, so as to know the attitude of teachers and parents towards the application service of college entrance examination through interviews.

$$\begin{bmatrix} T_{11}T_{12}\cdots T_{1n} \\ T_{21}T_{22}\cdots T_{2n} \\ \vdots \\ T_{n1}T_{n2}\cdots T_{nn} \end{bmatrix}\begin{bmatrix} S_{u1} \\ S_{u2} \\ \vdots \\ S_{un} \end{bmatrix} = \begin{bmatrix} R_{i1} \\ R_{i2} \\ \vdots \\ R_{in} \end{bmatrix}$$

Figure 2. Distributed computing for Mapreduce

For students of all the economically from the parents, parents' attitude for students ultimately whether or not a volunteer service is the most crucial role, through the data survey, most of the parents in the platform service

information accurately. In the high probability, it will accept the conditions of a certain paid service platform, and the teacher is also willing to accept a similar platform to help the students to the school, or through the teacher to the students volunteer to fill in suggestions.

According to the results of the survey, high school students and parents generally have the problem of asymmetric information about colleges and universities, and there is a considerable demand to accept the information provided by such services. Influencing the type and price of the final services may be related to factors such as economic location and service quality. The number of college entrance examination candidates in China is huge every year, so it can be seen that the potential market of college entrance examination application App is huge. In view of this kind of problem, facing a large number of college entrance examination students every year, its disadvantages are also obvious.

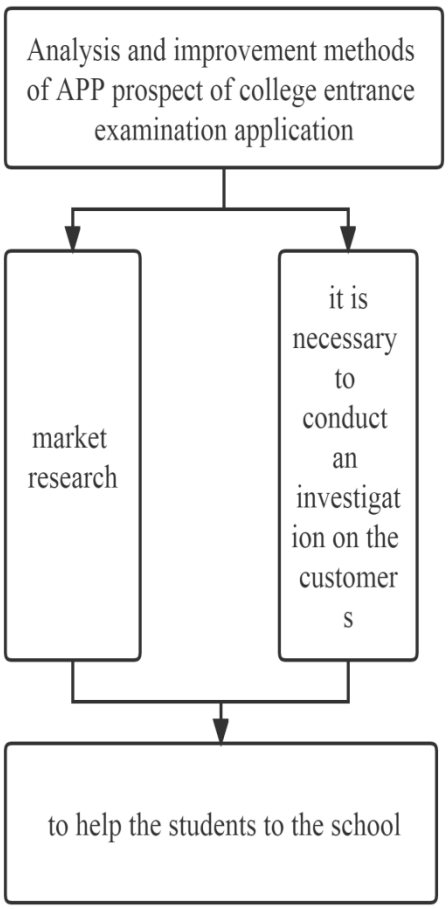

Figure 3.Analysis and improvement methods of APP prospect of college entrance examination application

The needs of each user will be different, which will cause great obstacles to the personalized service of college entrance examination volunteer filling. All kinds of objective information provided by App can also be obtained by customers with the help of big data [3]. Therefore, the technical method of voluntary application limited to objective information will hinder the quality of service. For example, ① College students connect with college entrance examination volunteer service groups, adopt many-to-one approach, directly provide users with the most direct and real college experience to give customers the correct guidance. ②Enterprise college evaluation system, that is, through enterprises for the number of college recruitment, recruitment majors and other directly let users observe the employment quality of colleges and universities. ③ Users directly connect to the system, for users, the user experience is the most realistic guidance. Also adopt the many-to-one approach, allowing multiple users to give feedback and suggestions to users, so that users can have a more direct feeling.

Table 3.The influence of individual needs on college entrance examination

The influence of individual needs on college entrance examination	
①	College students connect with college entrance examination volunteer service groups, adopt many-to-one approach, directly provide users with the most direct and real college experience to give customers the correct guidance
②	Enterprise college evaluation system, that is, through enterprises for the number of college recruitment, recruitment majors and other directly let users observe the employment quality of colleges and universities

③	Users directly connect to the system, for users, the user experience is the most realistic guidance. Also adopt the many-to-one approach, allowing multiple users to give feedback and suggestions to users, so that users can have a more direct feeling.

5 CONCLUSION

The voluntary filling is the most important part of the examinee after the college entrance examination. Choosing the right college and major has a great influence on the examinee's future development. Therefore, it is very important to obtain accurate information of the voluntary filling in the college entrance examination. With the advent of the information age and the emergence of big data technology, the application App for college entrance examination has emerged. Under the background of big data, people can solve the gap and asymmetry of their understanding of colleges and majors through apps such as college entrance examination application, and choose appropriate colleges and majors according to the information they have obtained and their interests and future development. In the process of the university entrance exam to college services, using the App to provide all kinds of information must be truthful and accurate, and pay attention to the legal problems in the process of information service. It can not transcend the bounds of the law, improve the way of providing volunteer service, not only can produce huge economic benefits, but also has great social significance, promote social fairness and harmony, and promote the development of socialism.

REFERENCES

[1] Cheng Lulu. Design and implementation of voluntary filling system based on multi-source heterogeneous data page rendering [D]. University of Chinese Academy of Sciences (Shenyang Institute of Computing Technology, Chinese Academy of Sciences), 2021.

[2] Du Yonggui. Application analysis of big Data Technology in Student Voluntary Application [J]. Software, 2021,42 (07): 178-180.

[3] Du irui. Analysis of the application of big data technology in the college entrance examination voluntary application [J]. China New Communications, 2019,21 (05): 36-37.

[4] Gu Jian. Research on key technologies of gaokao volunteer data analysis based on big data [D]. Changchun University of Science and Technology, 2017.

[5] Gao Lili, Fan Caihong. Design of the college entrance examination voluntary application and recommendation system based on big data [J]. Electronic Technology and Software Engineering, 2021 (22): 213-215.

[6] Hu Zuo Wei, Niu Xinbo, Liu Bo, Chen Ming, Du Huiyong. Design and implementation of the voluntary filling auxiliary system under the new College Entrance Examination reform [J]. Smart Computer and Application, 2019,9 (02): 175-179.

[7] He Bin. Design and implementation of the auxiliary decision system of college entrance examination score evaluation and voluntary filling [D]. Xidian University, 2006.

[8] Huang Daiqin, Zhou Qiang, Yu Feihong. Research on "Zhejiang New College Entrance Examination" based on big data analysis [J]. Computer Knowledge and Technology, 2019,15 (31): 75-78.

[9] Meng Zhen. The Design and Implementation of the College Entrance Examination Recommendation System Based on Spark [D]. Shandong Normal University, 2017.

[10] Meng Wei, Shang Yuhui, Zhao Xiaofei. "Artificial intelligence" helps the new college entrance examination volunteer filling inquiry [J]. University, 2021 (41): 140-142.

[11] Silver Hongyu. Design and implementation of the college Entrance examination volunteer recommendation system based on big data [D]. University of Electronic Science and Technology, 2018.

[12] Wan Jinjin, Hu Qianyu, Zhang Chunyan, Wei Weile, Zhao Xintao, Muhe Ta'er Reken. App of college entrance examination application under big data [J]. Markeweek, 2020 (01): 179-180.

[13] Yu Chao, Liu Guozhu. Voluntary filling algorithm based on big data analysis [J]. Journal of Qingdao University of Science and Technology (Natural Science Edition), 2020,41 (02): 113-118.