Big Data Mining Analysis of Key Indicators of Online New Energy Vehicles

Caidong Gu\textsuperscript{1*}, Zhaobin Liu\textsuperscript{2} and Ligang Fang\textsuperscript{2}

1 Jiangsu Modern Enterprise Informatization Application Support Software Engineering Technology Research and Development Center, Suzhou, Jiangsu, China
2 Suzhou Vocational University, International Education Park, 106 Zhineng Avenue, Suzhou, Jiangsu, China

Email: gucd@jssvc.edu.cn

Abstract. In this paper, more than 300 new energy vehicle key indicators of online sales are mined, and big data visualization analysis is carried out. In October 2018, big data mining analysis was conducted on the brand, model, pure electric endurance, slow charging time, fast charging time, battery capacity, maximum power, maximum torque, monthly sales volume, annual cumulative sales volume, monthly transaction price, user score, number of public praise, manufacturer guidance price (low) and manufacturer guidance price (high) of each new energy vehicle market in Sina auto.com. The results show that the average pure electric endurance of new energy vehicles is 210.3 km (with a maximum endurance of 613 km), and the best-known brand of new energy vehicles is BYD, with the highest market share of BYD and Toyota.

1. Introduction

New energy vehicles are the key development areas in the national strategic emerging industry development plan of the 13th five year plan. China’s production and sales of new energy vehicles rank first in the world for many years. In 2016, China’s production of new energy vehicles reached 517000, with sales of 507000. In 2016, BYD, Geely, BAIC and other enterprises entered the top 10 sales volume of new energy vehicles and passenger cars in the world, and domestic new energy buses were sold to more than 30 countries and regions in the world.

Can new energy vehicles achieve high-quality development and really become the breakthrough of China’s automobile industry? The independent R & D capability of core technology is of course the most important [1]. In this context, the significance of national monitoring and management platform for new energy vehicles, big data analysis and new energy vehicle evaluation index system based on this is self-evident. It can not only make the policy more accurate and effective, but also provide enterprises with really helpful high-quality services. New energy vehicles are developing rapidly with the application of big data, Internet of things and artificial intelligence (AI) technology [2]. Big data has become the hotspot of technology, and the era of big data has come. How to analyze and use huge data to promote business and production to produce value, it is necessary to formulate the development strategy of new energy automobile industry in advance [3]. At present, the research and analysis of big data of new energy vehicles is still in the exploration stage, and it is a new task to develop new energy vehicle development strategy based on big data.
It is very important for China to make a good strategy for the development of new energy vehicles. Based on big data analysis of new energy vehicle industry, through its big data technology and resource advantages, big data in public transport, logistics, sanitation, postal, power communication and other information, formulate countermeasures to make new energy vehicle become a new bright spot in China after high-speed rail and bridge, and play a leading role in demonstration. The new energy automobile industry, as the next tuyere industry, will promote the economy in all aspects. It will be a great help to the Rural Revitalization and the further deepening of the urban economy [4]. With the release of the national system reform dividend and the demand of environmental protection, the urban and rural new energy vehicle market is expected to usher in a wave of blowout development. Based on big data analysis, China’s new energy vehicle market structure optimization, urban-rural development balance, application industry layout, energy conservation and emission reduction, new energy supply, intelligent and intensive road traffic, etc. [5], can be developed in advance to promote the healthy and sustainable development of the industry [6].

2. Research Method Steps
We collect big data related to new energy vehicles from automobile professional websites, and obtain information based on big data visualization analysis software. We take Sina auto. Com as an example, in October 2018, it mainly collected about 300 kinds of key indicators of online vehicles, including brand, *model, pure electric endurance, slow charging time, fast charging time, battery capacity, maximum power, maximum torque, monthly sales volume, annual cumulative sales volume, monthly transaction price, user score, word-of-mouth number, manufacturer guidance price (low), manufacturer guidance price (high). Based on the analysis of these big data, the research results are obtained [7, 8].

We analyze and mining process of new energy vehicles: firstly, climb the mainstream Internet vehicle page, manage the planned or crawled URL through the URL management module, and then use the page download module to access and download the URL determined by the URL management module. Clean and save the downloaded content. Chinese uses hidden Markov chain as word segmentation, calculates TF-IDF scores to extract keywords, calculates Hamming distance, and makes comments on Chinese emotion classification. We Using the node, edge and degree of network graph, construct clustering coefficient, shortest path, number of points, small world network and so on. Spatial data is analyzed by using geographic information system, location-based services and latitude and longitude information of spatial data. Cloud storage and parallel computing are realized based on Hadoop, and data mining visualization is realized by tableau [9, 10].

3. Research Results and Analysis
We dig the big data of new energy vehicles obtained by Sina auto.com.

3.1. Average Value of Main Technical Indicators of New Energy Vehicles
As can be seen from figure 1, the maximum pure electric endurance is 613 km, and the maximum fast charging time is 0.5 h.

According to table 1 and figure 2, Tesla’s maximum endurance is 613 km, followed by BYD’s 500 km. The maximum endurance is still good.

According to table 2 and figure 3, the shortest fast charging time of GAC new energy and Chang’an automobile is 0.5 hours, followed by Jiangling group new energy and BAIC new energy by 0.66 and 0.6 hours. The minimum charging time of 0.5 hours is still too long for most users. If it is shortened to a few minutes, it is a better result.

3.2. Social Benefit Index of New Energy Vehicles
According to table 3 and figure 4, Toyota accounts for 19.8%, BYD accounts for 16.8% and BAIC new energy 11.1% are in the top three. BYD and BAIC new energy ranked in the top 10 in terms of their endurance mileage and fastest charging time, so they ranked in the top three in terms of market
share. Although Toyota did not rank in the top ten in terms of endurance mileage and fastest charging time, it was also in the top middle level, ranking first in terms of market share with its old brand and low price.

| Pure electric maximum endurance 613km | Charging slowest time 36h | Maximum power 1000kW |
|----------------------------------------|---------------------------|----------------------|
| Maximum battery capacity 100kWh       | Fastest charging time 0.5h| Maximum work 967N.m  |

**Figure 1.** The highest value of main technical indicators of new energy vehicles on Sina auto.com.

**Table 1.** Top 10 pure electric endurance distance of new energy vehicles on Sina automobile network (km).

| Endurance (km) | Tesla | BYD | Changan automobile | Jianghuai Automobile | BAIC new energy | Roewe | Know beans | Bao Jun | Zhongtai | GAC new energy |
|----------------|-------|-----|--------------------|----------------------|----------------|-------|------------|---------|----------|----------------|
| Maximum endurance | 613   | 500 | 300                | 400                  | 380            | 301   | 310        | 250     | 410      | 152           |
| Average endurance  | 589   | 400 | 300                | 251                  | 255            | 115   | 180        | 165     | 70       | 152           |
| Minimum endurance  | 565   | 360 | 251                | 170                  | 150            | 53    | 155        | 120     | 58       | 152           |
| Minimum endurance ranking | 1     | 2   | 3                  | 4                    | 5              | 6     | 7          | 8       | 9        | 10            |

**Note:** Tesla 特斯拉, BYD 比亚迪, Changan automobile 长安汽车, Jianghuai Automobile 江淮汽车, BAIC new energy 北汽新能源, Roewe 荣威, Know beans 知豆, Bao Jun 宝骏, Zhongtai 众泰, GAC new energy 广汽新能源.

**Figure 2.** Top 10 pure electric maximum endurance distance of new energy vehicles.
Table 2. Top 10 fast charging time of new energy vehicles on Sina auto.com (H).

| Quick charge (h) | GAC new energy | Changan automobile | Jiangling group new energy | BAIC new energy | Jianghuai Automobile | Roewe | BYD | Zhongtai | BMW | Tesla |
|------------------|----------------|-------------------|---------------------------|----------------|----------------------|-------|-----|---------|-----|-------|
| Minimum fast charge time | 0.5 | 0.5 | 0.66 | 0.6 | 1 | 0.5 | 1.1 | 0.5 | 2 | 4.5 |
| Average fast charge time | 0.5 | 0.65 | 0.66 | 0.875 | 1 | 0.67 | 1.5 | 1 | 3 | 4.5 |
| Maximum fast charge time | 0.5 | 0.8 | 0.66 | 1 | 1.5 | 3 | 3 | 6 | 3.8 | 4.5 |
| Ranking | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Figure 3. The fastest charging time of new energy vehicles on Sina auto.com.

Table 3. Top 10 market share of new energy vehicles on Sina Auto Net (%).

| Market (%) | Toyota | BYD | BAIC new energy | Roewe | LEXUS | Honda | Jianghuai Automobile | Zhongtai | Tesla | Bao Jun |
|------------|--------|-----|-----------------|-------|-------|-------|----------------------|----------|-------|---------|
| Market share | 19.8 | 16.8 | 11.1 | 9.1 | 5.8 | 4.8 | 4.5 | 2.3 | 2.3 | 2.2 |
| Ranking | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

From table 4 and figure 5 of the new energy automobile users’ word-of-mouth ranking of sina automobile network, BYD, Roewe and Lexus ranked in the top three.

Table 4. Top 10 users of new energy vehicles on Sina auto (%).

| Public praise | USER word of mouth | BYD | Roewe | LEXUS | Know beans | Song of vomit | TeslabMX | Zhongtai | BAIC new energy | Ferrari |
|---------------|-------------------|-----|-------|-------|------------|--------------|----------|----------|----------------|---------|
| User word of mouth | 732 | 157 | 160 | 154 | 149 | 133 | 86 | 83 | 80 | 79 |
| Ranking | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
3.3. Characteristics of New Energy Vehicle Market
Since 2015, the sales volume of new energy vehicles has ranked the first in the world for three consecutive years. In 2017, the growth rate, market share are the highest, and the technical level is significantly improved. In 2015, the sales of new energy vehicles reached 331000, a year-on-year increase of 3.4 times. In 2016, sales of new energy vehicles reached 507000, an increase of 53% year on year. In 2017, sales of new energy vehicles reached 777000, an increase of 53.3% year on year.

From table 5, we can see the three major trends of the development of new energy vehicles in China: first, the sales volume is growing rapidly, thanks to financial subsidies, license concessions, unlimited travel and other incentives. Since 2015, the sales volume has been the first in the world for three consecutive years, with a growth rate of more than 50%; in 2017, the individual consumption of new energy passenger vehicles exceeded 70%, indicating that the personal consumption market is rising rapidly; third, pure electric vehicles are still growing rapidly. It is the main force, accounting for 83.9% in 2017, higher than 80.7% in 2016 and 74.6% in 2015.
### Table 5. Sales of new energy vehicles in China in recent three years.

| Project                      | 2015 | 2016 | 2017 |
|------------------------------|------|------|------|
| Sales volume of new energy vehicles (10000 units) | 33.1 | 50.7 | 77.7 |
| Proportion of pure electric vehicles in new energy vehicles | 74.6% | 80.7% | 83.9% |

#### 3.3.1. New Energy Vehicle Related Big Data Collection and Sorting

New energy vehicle big data mining analysis, big data application mode, big data standardization research. We collect and analyze the big data of new energy vehicle technology, performance parameter configuration, quality, market, customer, after-sales service, supporting facilities and future development trend, so as to provide the next development direction for the enterprise. The big data of user consumption tells you what brand of new energy vehicles you want to buy, the big data of manufacturers interprets the configuration level of new energy vehicles, the big data of dealers interprets the characteristics of Regional Automobile sales market and automobile consumers, the big data of network access shows that automobile users access the network resources with high activity, and provides data with certain reference value for automobile marketing. For government policy big data collection, such as limited license plate, limited platoon, limited purchase and other policies, it also makes new energy vehicles become the key development, focus object and direction sign of manufacturers in the future. In the key research fields of made in China 2025, new energy vehicles and intelligent Internet connected vehicles are also taken as the main development direction, among which low-carbon, information and intelligent vehicles are taken as the core technology strategy. Smart car based on big data is an important strategic direction for the transformation, upgrading and sustainable development of China’s automobile industry. In this way, the positioning of new energy vehicles is determined.

#### 3.3.2. Use Big Data to Analyze the Development Strategy of New Energy Vehicles

Use the big data of new energy vehicle configuration of manufacturers to analyze the current situation of automobile technology, develop technology upgrading strategies, use the big data of new energy vehicle sales of dealers to develop after-sales strategies, develop R & D direction based on the hobby of new energy vehicle sales crowd, develop regional products and market strategies based on the big data of new energy vehicles, and focus on buses, taxis, car sharing, logistics vehicles, etc. For further popularization and promotion of new energy vehicle big data in the field, formulate corresponding policies and measures. So that new energy vehicles constitute a relatively complete set of R & D, design, production, sales, promotion and application, development planning strategy system.

### 4. Results and Discussion

(1) Based on big data mining, this paper obtains the average data and the maximum and minimum data of more than 300 new energy vehicles currently sold in real time for the first time. For example, the average pure electric endurance of new energy vehicles is 210.3 km (the maximum endurance is 613km), and the average fast charging time is 1.7 hours (the minimum time is half an hour).

(2) Based on big data mining, this paper obtains more than 300 major social and market indicators of new energy vehicles currently sold in real time for the first time. For example, BYD has the best reputation for new energy vehicles, and BYD and Toyota have the highest market share. There will be a growing demand for new energy vehicles. The national, local and people’s attention to environmental protection, the continuous subsidy policy for new energy vehicles, the continuous improvement of the quality of new energy vehicles, and the more convenient and popular use.

(3) Buyers need new energy vehicles with good quality and low price, safety and comfort, long endurance, convenient charging and short time to calculate in minutes. The automobile industry is trying to develop high-quality new energy vehicles with stronger endurance, shorter charging, safety and intelligence. For example, Li Xiang’s “ideal smart manufacturing” brand is an electric vehicle without endurance anxiety (with an endurance of 1000km). Several new energy vehicles, such as BYD, can charge only half an hour at the fastest. With the breakthrough of Internet of things, Internet
of vehicles and automatic driving technology, the elimination and upgrading of oil and gas vehicles, the development momentum of new energy vehicles will be stronger.

(4) Based on the power, solar energy, wind power, nuclear energy and other power driven new energy vehicles gradually show the trend of diversified development. New energy vehicles based on electric energy have become the main force of new energy vehicles. In the future, with the safe use of photovoltaic power generation, wind power generation and nuclear energy, the development of new energy vehicles based on renewable future energy will be diversified.

(5) At present, China’s new energy vehicles still have insufficient core technology control ability and imperfect use measures. The convergence and layout of national and local policies are very important. The core technology of new energy vehicles is hollow, and there is no top-ranking enterprise like Huawei, which is very important for China’s ambition of using new energy vehicles to achieve corner overtaking in the automotive field. It is very important for first-class enterprises to work together, and more importantly, some enterprises with good development momentum in the future should be merged and reorganized to jointly build a new aircraft carrier of new energy vehicles in China.

In the future, the state and local governments should strengthen the overall planning and policy convergence in response to these problems, and work together to promote the rapid implementation and solution of imperfect problems.

Acknowledgments
This work is supported in part by NSFC under Grant No. 61672372 and 61472268.

References
[1] Xian X F, Zhao P P, Sheng Victor S, Fang L G, Gu C D, Yang Y F and Cui Z M 2016 Stratification-based outlier detection over the deep web Computational Intelligence and Neuroscience.
[2] Chen R Y, Wang Q L and Xu W 2019 Mining user requirements to facilitate mobile app quality upgrades with big data Electronic Commerce Research and Applications.
[3] Guo C H 2019 Big data analytics in healthcare: data-driven methods for typical treatment pattern mining Journal of Systems Science and Systems Engineering.
[4] Liang Y H and Wu S Y 2015 A scalable and effective frequent itemset mining algorithm for big data based on MapReduce framework IEEE International Congress on Big Data.
[5] Gu C D 2017 Fast discrepancy identification for RFID-enabled IoT networks IEEE Access.
[6] Gong W, Liu J and Yang Z 2017 Efficient unknown tag detection in largescale RFID systems with unreliable channels IEEE/ACM Transactions on Networking 25 (4) 2528-2539.
[7] Gong W, Stojmenovic I, Nayak A, Liu K and Liu H 2016 Fast and scalable counterfeits estimation for large-scale RFID systems IEEE/ACM Transactions on Networking 24 (2) 1052-1064.
[8] Gu C D, Gong W and Nayak A 2016 Identifying discrepant tags in RFID enabled supply chains Proceedings of WASA.
[9] Beng C, Kian L T and Tung A 2010 Sense the physical, walkthrough the virtual, manage the co (existing) spaces ACM SIGMOD Record No 3.
[10] Zheng Y, Zhang L, Xie X and Ma W Y 2009 Mining interesting locations and travel sequences from GPS trajectories Proceedings of the 18th International Conference on World Wide Web.