

# Innovative Research on Market Analysis of New Energy Cars Based on Online Reviews

Bin Li, Xiaofang Liu

China Auto Information Technology (Tianjin) Co., Ltd, China

---

**Abstract:** With twitch the intensification of competition in the new energy car market, it is particularly important for car enterprises to prepare to grasp the market competition situation and changes. In this paper, by extracting the model comparison information of consumers in online reviews, the competition situation among models is shown through the directed attribute relationship diagram, which is used to point out the shortcomings of the existing research methods of new energy car market competition situation, and propose a new market analysis idea. The results show that when consumers compare models, they break the original market segmentation, cross level and cross field comparison, and there are significant differences in the range of competitive products at different time points, which will help improve the existing market competition situation research methods in the process of serving automobile enterprises in the future.

**Keywords:** Online comments, Competitive situation, New-energy car, Directed attribute relationship.

---

## 1. Introduction

With the improvement of people's material level, cars have changed from a luxury item in the past to a necessary household item for travel, thus stimulating the booming development of the automobile industry. By 2022, there are more than 1,000 passenger car models on sale. However, too many models lead to intensified competition in various market segments, which not only affects the sales of automobile companies, but also shortens the life cycle of each product. In this context, accurately grasping the competition situation of market segments helps auto companies to clarify product positioning, better predict sales performance, and then rationally plan production and marketing plans. Therefore, accurate analysis of market segments is becoming more and more important for car companies.

## 2. Research Background

### 2.1. Research on Competitive Situation of New Energy Vehicles

Based on design and technical characteristics, new energy vehicles mainly refer to cars used to carry passengers and their luggage, with no more than nine seats including the driver. Based on the classification widely recognized by consumers in China's auto market, new energy vehicles can be classified as basic passenger vehicles (sedans), multi-purpose vehicles (MPVS) and sport utility vehicles (SUVs). Based on the size, battery capacity and other factors, it can be further subdivided into small, medium and large cars. Among the models sold, the car market is the main one, which contains the most models and the most intense competition.

Based on national standards, it is mainly classified on the basis of the appearance and function of new energy vehicles. But the actual car companies, businesses and car consumers pay more attention to the classification according to the level of vehicles. The grade of a car includes not only the price, but also a large number of technical parameters, such as length, wheelbase, displacement, etc., which has become the classification method of the current mainstream market segment. On the other hand, according to their own marketing

purposes, car companies will also subjectively classify different grades, crowned by mass production, high-end, luxury and other classification standards, with a strong subjective color, and then cater to the vanity of some car consumers. In addition, some scholars add concepts such as value based on the market value performance of new energy vehicles and brands, hoping to further refine the existing market segments through the combination of sales volume, price and value. [1]

The reason why some scholars introduce the concept of value is mainly because automobile consumers will refer to some dynamic information in the process of car purchase, such as brand premium, word-of-mouth, reviews and other factors, which will affect the final decision of car purchase. However, the existing research methods are mainly carried out through questionnaire survey.

### 2.2. Review and Research

The rise of review research is mainly due to the progress of computer technology, which makes people's communication break the limitation of time and region, and can communicate on the Internet anytime and anywhere. Car consumers no longer listen to the propaganda and expert recommendations of car companies, but more widely refer to the information shared by other car owners, and actively reply and exchange. According to the latest conclusions in marketing, the current marketing model is gradually changing from traditional AIDMA (Attention Interest Desire Memory Action) to AISAS (Attention Interest Search Action Share) mode shift. In the new marketing rules, the emergence of two "s" with network characteristics -- search and share, points out the importance of search and sharing in the Internet era, rather than blindly instill one-way ideas to users, which fully reflects the influence and change of the Internet on people's lifestyle and consumption behavior. [2] Based on this, a large number of scholars began to focus on text mining technology to extract valuable information in reviews.

Text Mining is a process of obtaining interesting or useful patterns from unstructured text information. In the late 1950s, H.P.Luhn carried out pioneering research in this field and put forward the idea of word frequency statistics for automatic

classification. Subsequently, many scholars have carried out effective research on the automation and recognition accuracy of text mining algorithms, and developed algorithms such as association rule extraction, semantic relationship mining, text clustering and topic analysis, trend analysis and so on. At present, supervised and unsupervised machine learning algorithms are the main focus. [3] Domestic research started late, but with the promotion of Ali, Tencent, Baidu, IFlytek and other domestic first-class big data companies, remarkable achievements have been made. At present, our country's text mining is still in the training of digesting and understanding foreign related theories and technologies, and matching Chinese related algorithms.

Based on the existing research direction of competition situation in the new energy vehicle market, this paper tries to propose a solution through review research to quickly identify the competitive relationship between models that is closest to consumers' cognition.

### 3. Research Methods

#### 3.1. Jieba word segmentation thesaurus

Jieba word segmentation is selected as the basic library to extract useful information from the review information center. Jieba word segmentation is an excellent third-party Chinese library of python. Its principle is to construct a prefix dictionary to scan the word graph based on the statistical dictionary, and then use the prefix dictionary to segment the input sentence, and get all the segmentation possibilities. Its modes can be divided into full mode, precise mode and search engine mode.

Since this paper studies the competitive relationship between car models, the default precise mode can be selected, and a car alias library is constructed as an additional thesaurus to identify high-frequency words.

#### 3.2. Synonyms Kit

Synonyms is a Chinese synonym toolkit, which is mainly used for text alignment, recommendation algorithm, similarity calculation, semantic offset, keyword extraction, concept extraction, automatic summarization, etc. Since consumers will call parts by nicknames or aliasing in the process of communication, it is difficult to establish a more complete alias library in the form of manual marking, thus affecting the lack of samples of useful reviews. Therefore, based on jieba lexicon, the Synonyms toolkit is added to identify the aliasing of supplementary car models. The basic technology used in the toolkit is Word2vec, which uses a large amount of data and is trained with context information to map words to low-dimensional space and generate relationships, which can be quantified as distances and can be further retrieved using the distances between words.

Synonyms refers to 200 million trained word vectors from Wikipedia by default. Jieba segmentation lexicon is introduced to replace the existing Synonyms lexicon, which can be closer to the professional terms of the automobile industry, and then obtain the relevant aliasing of car models.

### 3.3. Directed attribute diagram analysis

In the process of consumers' actual car selection, there is not necessarily a strong comparison relationship between car models. However, the existing competition situation ignores the strength of the competitive relationship between models, which makes the division of competitive product circle doped with certain subjective factors and affects the accuracy of the results.

Referring to the banking industry, the capital exchange relationship (fixed trading circle) is often locked through the strength of the lending relationship between individuals, and the directed attribute relationship graph is introduced to identify the strength of the contrast between models. [4] The problem of distance between models is not solved, and the Ising ferromagnetic model algorithm is introduced to solve the problem. The Ising ferromagnetic model is a spatial random field in which each atom is subject to two forces. One is the force of the surrounding atoms and the force of the external field, and the other is the disturbance of the environmental noise to produce random deflection, that is, the direction of each atom is affected by these two forces. Suppose we use variables to represent the orientation of each atom, as shown in Figure 1.

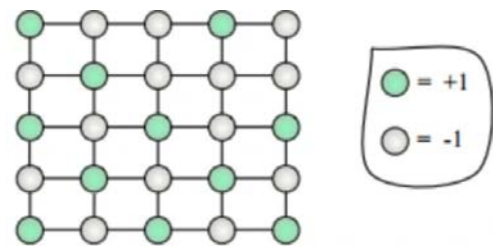


Figure 1. An example of the Ising ferromagnetic model

## 4. Research Results

### 4.1. Research Scope

This paper takes the Autohome portal as an example and selects about 23,000 comments from January 1, 2020 to August 31, 2021.

PYTHON is used as the word segmentation and training tool, and PBM platform is used to complete the visual display of competition situation and the calculation of vehicle type relationship.

Supplement the dimension fields of existing market segment classification, power type, price (mean), sales volume and model level of relevant models.

### 4.2. Research results

By analyzing and identifying 3958 valid samples, 951 models were counted and 1741 aliases were built. Among them, MPV involves 118 models, 494 valid samples, 1572 concerns, and 17 related market segments; SUV involves 574 models, 3375 valid samples, 30878 concerns, and 20 related market segments; There are 603 car models, 5078 valid samples, 54985 concerns and 17 related market segments. As shown in Table 1.

Table 1. Statistical result

Grade of product	vehicle type	Number of comparisons	Attention	market segment
MPV	118	494	1572	17
SUV	574	3375	30878	20
Sedan	603	5078	54985	17

All samples were imported through PBM platform to generate the competition diagram, as shown in Figure 2.

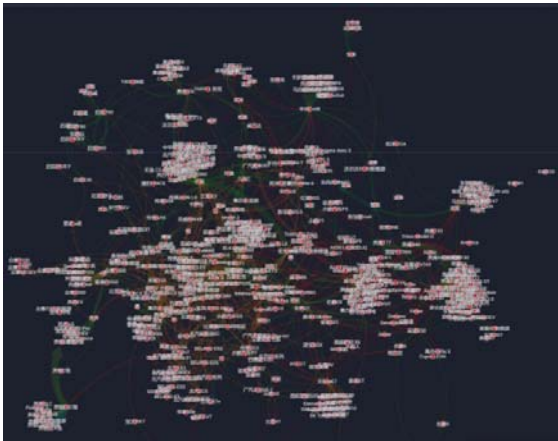


Figure 2. Global diagram of competitive relations

The dot in the figure represents a model, the line represents the mention relationship between the models, the direction of the arrow represents the direction in which the owner of the product or the potential owner of the product actively mentions the competing product, the color of the line represents the market segment, and the thickness of the line represents the number of mentions.

Due to the mention relationship and the number of mentions between competing products, the spacing distribution between models is uneven, and the classification of models with strong contrast relationship is obvious, forming a distinct competition situation, as shown in Figure 3.



Figure 3. Schematic diagram of competitive product circle

Taking Trumpchi AION as an example, in 2020, it formed its own product system with GS4 and AION V as cross-competing products, mainly related to Xpeng and Guangfeng models, involving pure electric, hybrid and gasoline vehicles. By 2021, the product system with Trumpchi M8 as cross-competing products will be formed, and Honda's competing products will be added, and the competition relationship with XPeng will be weakened. It can be seen that in the actual car selection process, consumers' choice of competing product circle is not fixed, but changes with model iteration and time, as shown in Figure 4.



Figure 4. Schematic diagram of Toyota C-HR2020&2021 competitive products

Combined with the driver, price and other factors, can further lock the relevant competitive product models, price belt and other information. Taking a car model as an example, Asialong, Lingshang and Corolla are identified as core competitive products (three cars with more two-way comparison times) through the competition situation diagram. Through the quadrant diagram of car length and price combination, the middle comparable cross-competing products can be selected. Through the comparison of model price bands, the price range in users' mind is locked, as shown in Figure 5.

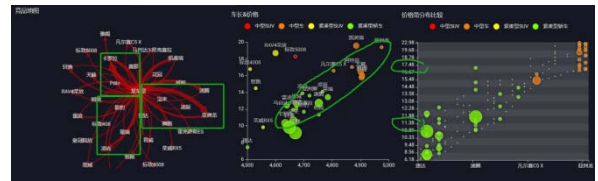


Figure 5. Schematic diagram of competitive product analysis process of a car model

## 5. Conclusion

It is found that the comparison information of car models can be extracted by extracting the user's review information. On the other hand, the directed attribute relationship graph can better pull in and away from the core competitive product models, so that the competitive relationship can be more intuitive display. At the same time, combined with price, car length, time and other factors, we can understand the changes in consumers' choice of competing products, which will help us better understand the reference factors for consumers to judge competing products in the future.

This study only verifies that the competitive relationship between models can be identified through useful reviews. In the future, empirical research can be conducted to quantify the reference factors that affect consumers' car choice, further explain the competition situation formed, and provide necessary reference basis for automobile companies to better design model positioning.

## References

- [1] WANG X J. Empirical Analysis on the Competition Situation of Passenger Car Market Based on Panel Data Model [D]. Beijing University of Technology, (2017).
- [2] ZHU M. Analysis of SNS Marketing rules and applications under the condition of Web2.0 [J]. Modern Marketing (Academic Edition),2011(07):75-77.
- [3] QIAN X. Research on the Construction Method of Enterprise Technology Innovation Index System Based on text Mining [D]. Beijing jiaotong university, 2021. DOI: 10.26944 /, dc nki. Gbfju. 2021.000368.

- [4] Chen Y. Research on user directed relationship strength based on Feature attribute, network structure and social behavior [D]. Zhejiang Gongshang University,2018.
- [5] Zhu Zhangqi. One-dimensional long-range ferromagnetic ising model critical index research [D]. Nanjing university of aeronautics and astronautics, 2019. The DOI: 10.27239 /, dc nki. Gnhhu. 2019.001013.