

From Coarse to Fine: A Bias Calibrated, Attribute Level Sentiment Analysis Framework with Dynamic Monitoring for Automotive Export Value Perception

Huiyuan Yang^{1*†}, Jingjing Yan^{2†}

¹American Heritage Schools, Plantation, USA

²Shandong Experimental High School, Jinan, China

*Corresponding Author. Email: pl1003685@ahschool.com

†These authors contributed equally to this work and should be considered as co-first authors

Abstract. China has developed as one of the world's largest automobile exporters, yet export prices have declined despite the increased volume, suggesting that growth in exports is still predominantly influenced by cost rather than value. This study introduces a bias-calibrated framework along with dynamic temporal monitoring to examine consumer value perceptions across six main characteristics. These include battery, charging, after-sales, climate, road, and, namely, price. After analyzing 8029 Reddit comments and aligning them with monthly export prices from CAAM, we discovered the highest sentiment (0.30) and mention rate (18.6%) came from prices. On the other hand, after-sales had the lowest score (0.16). Temporal analysis highlighted the seasonal cycles along with the declines. Spearman correlation revealed that battery sentiment positively correlated with export prices, with a ρ value of 0.38, which indicates that value directly leads to pricing for this technology. Price sentiment shows no correlation, which supports our hypothesis of the cost-driven logic. After-sales sentiment correlates negatively with a ρ value of -0.31, which suggests that price compensates for deficiencies in service. These findings suggest that it would be beneficial for China to adopt a hybrid pricing strategy for more exaggerated outcomes.

Keywords: Electric vehicles, China automobile exports, Sentiment analysis, Export pricing, Consumer perception

1. Introduction

The global automotive industry is currently being reshaped by electrification and shifting trade patterns [1]. China has emerged as the world's largest automobile exporter, with volumes rising from 4.91million units in 2023 to an estimated 8.07 million units in 2025, while the average export price fell from \$19,000 to \$16,000 over the same period [2]. This trend of a rise in volume and decline in price indicates that export growth remains primarily driven by cost rather than value-based [3]. In order to stay competitive in an international market, Chinese automobile corporations thus require a strategic shift toward consumer perception of product value, demanding a cross-regional analysis of product performance, infrastructure reliability, and service quality.

Past research on electric vehicle user satisfaction has grown alongside the methods of analysis. Early work utilized term frequency measures to identify adaptation barriers that may come with these vehicles, such as range and price sensitivity [4]. As time progressed, researchers started using embedding-based models, finding that overall sentiment was consistent. LDA topic modeling discovered that there was similarity between brand image, environmental concerns, and the performance of the vehicle [5]. More recently, Liu et al. [6] analyzed the reviews of charging stations with ChatGLM and discovered that there was a negative relationship between user ratings and station usage, which they attributed to the congestion within these areas rather than the physical problems associated with these stations [7]

Even with these advances, there are four underlying problems remaining. First, sentiment analysis persists in being broad, labeling whole documents as positive or negative without separating specific areas like battery durability or climate tolerance. Second, topic modeling assumes that topics stay unchanged over time and does not indicate shifts in consumer demand. Third, cross-cultural application is widely limited despite the fact that the market is global. Lastly, most studies do not account for platform-specific biases and they rarely connect sentiment findings to export pricing strategy. Due to these gaps, existing research does not adequately adopt NLP as a tool in clear guidance for export pricing decisions.

In order to lessen this limitation, this study proposes a deviation-correction emotional analysis framework that integrates dynamic time monitoring functions. The mechanism corrects the platform deviation by weighting individual feedback, and sorts the emotional score according to six different dimensions and dynamically tracks them over time. In addition, the method uses Spearman rank correlation to analyze the correlation between these scores and export price data, so that emotional responses can be accurately evaluated based on consumers' values and interests.

2. Related work

Traditional research from the perspective of electric vehicle users has evolved from simple word frequency analysis to more advanced natural language processing methods, such as classification technology and large-scale language models based on Transformers. Although these studies have laid the foundation for understanding consumer concerns, there are still technical limitations.

2.1. Coarse-grained emotional analysis

Traditional research methods often simplify complex consumer feedback into a single continuous value, thus summarizing opinions of different natures (such as various complaints) into unified quantitative indicators. For example, [4] used word frequency analysis to convert the entire forum post into binary indicators such as "duration anxiety" and "price sensitivity", and finally classified them into a single category [8]. [advanced word2vec-based CNN classifiers, yet they still reduced charging station reviews to positive/negative sentiment classifications. Liu et al. [6] sought to use ChatGLM to analyze 948 U.S charging stations, but they maintained a restricted classification of negative, neutral, and positive. This coarse output discards both intensity of information; a mildly negative comment such as "range could be tier" is indifferent from a strongly negative comment about "battery failed after just six months", and specified dimensions, conflating complaints about battery degradation with complaints about charging speed.

2.2. Topic modeling assumes static structures

Current approaches treat topic distributions as one incapable of capturing how consumer discourse evolves in response to events such as policy shocks, changes in technology, or seasonal cycles [5, 9, 10]. Suresha and Tawari applied LDA to 45,000 EV-related comments and identified static themes across a fixed time window. Bhatnagar and Choubey performed a similar study with LDA to extract topic classification on charging station reviews, but their analysis was limited to a cross-sectional analysis. These approaches treat topic distributions as undisturbed by time, so consequently, it results in capturing how consumer discourse evolves in response to policy interventions. (e.g the 2024 EU tariff announcement), technological shocks (e.g lithium price fluctuations), or seasonal cycles.

2.3. Platform specific biases remain uncalibrated

User generated content platforms contain sampling biases with upvote rate, yet existing work treats comments as i.i.d. samples without correcting for collider bias. Liu et al. [6] identified a counterintuitive negative correlation between user ratings and charging station utilization, mainly due to the congestion at charging sites. However, they did not use calibration to attempt to correct this bias. Borlaug et al. recorded geographic and temporal variation in charging station utilization but similarly did not adjust for the sampling bias found in user-generated content platforms, where comment visibility is controlled by upvote counts. Standard practice treats each comment as an independent and identically distributed sample [6, 8], ignoring that observed data are conditioned on a stochastic ranking and feedback process, which poses as a violation of i.i.d. assumptions that induces collider bias and distorts estimated relationships between sentiment and market outcomes.

3. Analytical framework

3.1. Overview of analytical framework

This study uses a multi-stage analytical approach that transforms unstructured user generated content into indicators of consumer value perception. The approach consists of four successive stages:

As shown in figure 1, the framework comprises the following stages: data processing, sentiment analysis, temporal aggregation, and statistical correlation analysis with export pricing. The framework demonstrates how raw Reddit comments can be transformed into consumer value metrics and linked to export pricing outcomes.

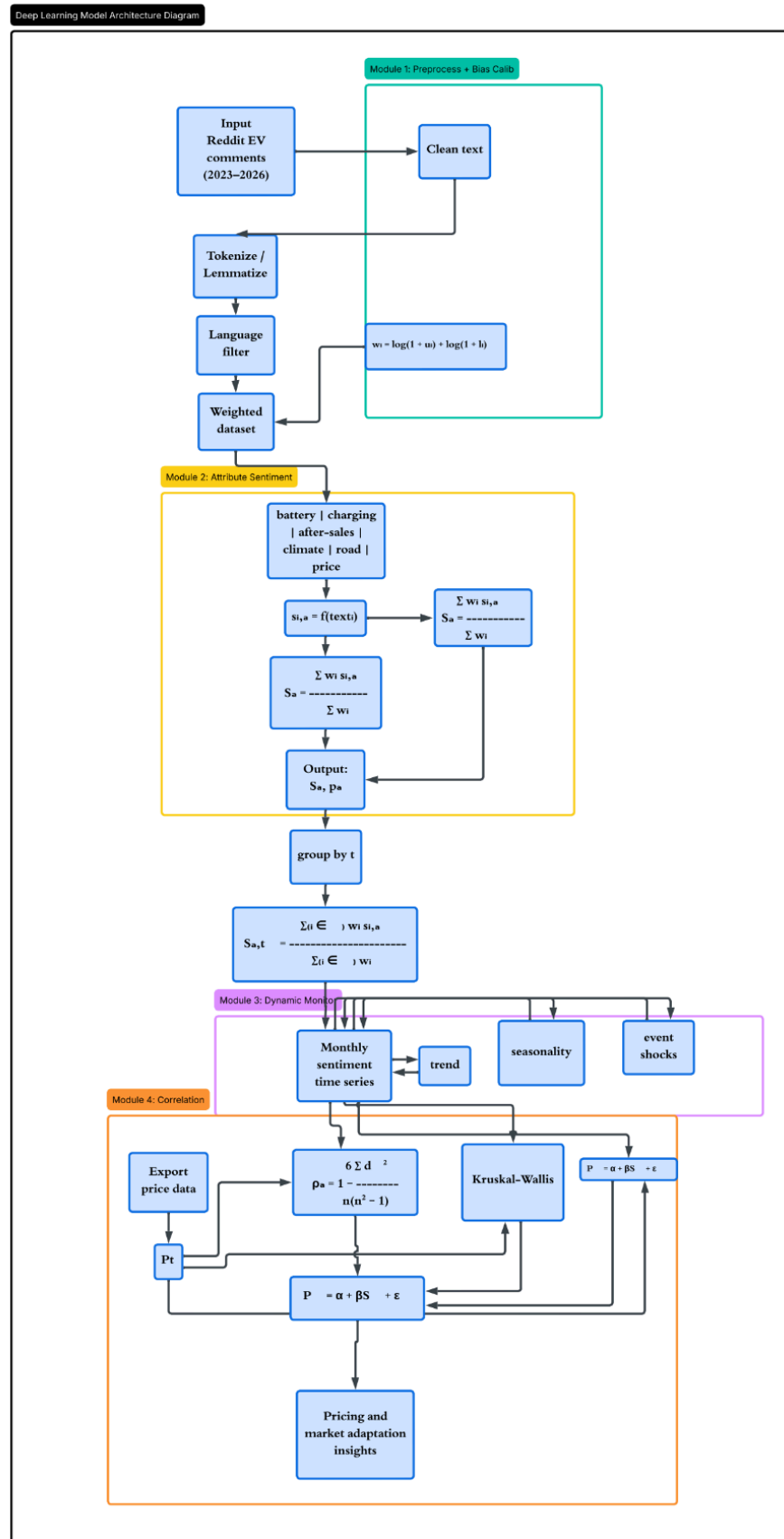


Figure 1. Overall framework of the proposed sentiment analysis methodology

(1) data processing and bias calibration, (2) extraction of sentiment at the attribute level, (3) temporal aggregation to monitor changes, and (4) statistical correlation with export prices. Each stage is designed to address specific limitations of previous studies, particularly too broad sentiment scales and insufficient calibration, the latter being the focus of this study

3.2. Data processing and bias calibration

3.2.1. Data collection

The data on car product reviews were sourced from Reddit, a social news and discussion platform that is regarded as a valuable resource for academic research [6]. This study uses the official Reddit API and collects comments through the Python Reddit API Wrapper (PRAW) tool. The data covers seven car-related sub-sections.

In order to focus on Chinese car brands for the export market, this study adopts a keyword-based search method. The selected keywords include specific terms such as "BYD Atto 3 review" and "MG4 EV long term review", as well as "Chinese Electric car quality" (China's electric car quality) and other broader expressions. For each keyword, up to eight relevant discussion posts were grabbed and all the comments from these posts. The final corpus contains 8,029 English comments, spanning from January 2023 to March 2026.

3.2.2. Text cleaning and normalization

Raw comments obtained from Reddit have been systematically cleaned. Regular expressions are used to eliminate elements that are not necessary for research, including but not limited to HTML tags, hyperlinks and special characters. All texts are converted to lowercase to ensure consistent format. Use the abbreviation expansion library to restore the abbreviation form to a complete form, so as to unify the lexical expression (for example, convert "don't" into "do not"). Comments with less than 15 characters after cleaning are eliminated because the amount of information contained in these comments may not be enough to support meaningful analysis.

Only the English comments were retained from the comments, as the CADER sentiment analyzer analyzed in this study is optimized when used for English language social media content.

3.2.3. Bias calibration through comment weighing

A key difference and innovation in methodology is the introduction of comment-level weights to adjust for platform specific biases. Prior research by Liu et al. [11] demonstrated that high usage of charging stations has a tendency to receive a disproportionate amount of negative reviews. This is a phenomenon attributable to the congestion effects rather than a perceivable issue of charging station quality deterioration. Frequently, user generated content platforms demonstrate a systematic bias where the more upvoted content tends to carry greater community validation and therefore is more prevalent throughout the platform.

To mitigate such biases mentioned, each comment is assigned a weight w_i calculated as:

$$w_i = \log(1 + u_i) + \log(1 + l_i) \quad (1)$$

The logarithmic equation compresses the extreme values, and as a result prevents a small number of highly upvoted comments from dominating the analysis. u_i is defined by the number of upvotes received by comment u_i . l_i indicates the text length measured in characters. This weighting scheme

ensures that the corresponding sentiment aggregations reflect both the volume of discussion and also the intensity of engagement within the community.

3.3. Attribute level sentiment analysis

3.3.1. Definition of value attributes

To move beyond the coarse positive/negative classification, six attributes were assigned that correspond to each dimension of consumer value perception on Chinese electric automotive (Table 1). These attributes were derived from a synthesis of previous research.

Table 1. Attribute of definitions and keyword sets

Attribute	Description	Example Keywords
battery	Battery performance, range, degradation	battery, range, mileage, degradation
charging	Charging speed, infrastructure availability	charge, charging, charger, fast charge
after-sales	Service, maintenance, warranty, repairs	service, warranty, repair, maintenance
climate	Performance in cold/hot weather conditions	cold, winter, snow, hot, summer, heat
road	Driving experience, handling, road conditions	drive, handling, road, suspension, narrow
price_value	Price perception, cost, value for money	price, cost, value, expensive, cheap

Each attribute is associated with a set of keywords designed to indicate the number of mentions within that category of user comment.

3.3.2. Sentiment scoring with VADER

The output form of sentiment analysis has changed from a single dimension to a comprehensive mode covering multiple attributes. For each attribute mentioned, the system will use prompt engineering to record three types of information: continuous emotional score, five levels of emotional intensity (extremely negative, negative, neutral, positive, extremely positive) and for strong emotions (score below 0.1 or above 0.6) comments. In addition, the optimized large language model will also extract problem attribution. This matrix form converts each comment into a structured feature description, thus ensuring the relevance of the evaluation while retaining the evaluation location information.

3.3.3. Weighted attribute aggregation

The weighted summary provides an overall view of the statistics. For each attribute a (2), the weighted average emotion is:

$$\bar{s}_a^{weighted} = \frac{\sum_{i \in M_a} w_i s_{i,a}}{\sum_{i \in M_a} w_i} \quad (2)$$

M_a represents the set of comments that mentions attribute a . The mention rate is identified as $pa=|M_a|/N$, as N is the total comments. These metrics allow for the possibility of comparing of both consumer attention and evaluation.

3.4. Temporal aggregation for dynamic monitoring

3.4.1. Monthly time series construction

At this stage, the evolution of various emotional indicators is evaluated by constructing monthly time series, so as to overcome the limitations of traditional theme modeling. The comments are grouped by month according to the timestamp, which leads to a total of 39 observation months from January 2023 to March 2026. For each month t and attribute a , the weighted average emotional value is calculated as follows:

$$\{s\}_a^{weighted} = \left(\frac{\sum_{i \in M_a} w_i s_{i,a}}{\sum_{i \in M_a} w_i} \right) \quad (3)$$

where, $M_{a,t}$ indicates any comment that mentions attribute a in t month. Thus, a time series is generated for each attribute, which allows us to capture its fluctuations over a long period of time.

The time series generated by this reveals three completely different patterns: long-term trends reflect cross-year emotional changes and the evolution of consumer cognition; the seasonal cycle captures the fluctuations of winter months and consumers' reactions to these laws; and the shocks associated with specific events identify policies adjusted or sudden emotional changes caused by external events such as tax implementation. This comprehensive approach overcomes the limitations of the traditional LDA model, because traditional models often regard subject distribution as an element unrelated to time.

3.5. Statistical association with export pricing

3.5.1. Theoretical framework

The final stage of this study aims to test whether consumer value perception (measured by multi-attribute emotional indicators) affects the export pricing of Chinese electric vehicles, and at the same time analyze its correlation with actual export prices. The export price data comes from the China Association of Automobile Manufacturers (CAAM), which matches the monthly average export price data of passenger cars (HS code 8703) provided by the association.

This study explores two completely distinct hypotheses. Under the "cost-driven" assumption, export prices are mainly determined by production costs and competitive conditions. At this time, consumers' value perception depends on other factors; according to this hypothesis, there should be no systematic correlation between emotional indicators and prices. On the other hand, under the "value-driven" assumption, consumers' value perception has a significant impact on pricing rights; this means that when consumer evaluation is positive, attributes such as battery performance or after-sales service should be positively related to export prices.

3.5.2. Spearman rank correlation analysis

A Spearman rank correlation was adopted as the main analysis tool. This indicator is used to evaluate the monotonic relationship between variables, and there is no need to assume that there is a

normal or linear relationship between the factors. Given that emotional data is often unevenly distributed, Spearman rank correlation is chosen due to its ability to measure the relationship between variables without assuming normal distribution; this characteristic allows it to be suitable for analyzing emotional time series. For each attribute a , the calculation formula of its correlation coefficient is as follows:

$$\rho_a = Spearman\left(\bar{s}_{a,t}^{weighted}, P_t\right) \quad (4)$$

3.5.3. Robustness checks and multi-method validation

In order to address various potential problems, this study adopts a variety of verification methods. The first of which being the Kruskal-Wallis test, which is used to examine whether there are significant differences in attribute emotions between different Subreddit communities. This helps to deeply understand whether the observed correlation comes from deviations in a specific community or varies with the context. In order to measure the joint relationship between attribute emotions and export prices, the study adopts a multivariate linear regression (OLS) model to ensure that all relevant variables are taken into consideration. However, given that social media data often does not conform to the normal distribution assumption, the interpretation of the results of the OLS model needs to be cautious. This study combines multiple analytical methods to minimize the risk that the observed relationship is only an outcome of a specific modeling hypothesis, thus providing stronger evidence for the relationship between consumer value perception and export pricing.

4. Experimental results and analysis

4.1. Datasets

This study constructs a comprehensive data set by integrating official export statistics and user-generated content from online car communities (such as Reddit). When selecting these data sets, this study aims to reveal and narrow the gap between the actual situation of the international market and consumer perception; this is crucial to verify the two assumptions of "cost-driven" and "value-driven".

4.1.1. Export trade data

The data set used in this study was extracted from the official statistical report of the China Association of Automobile Manufacturers [2], which covers the export data of new energy vehicles from January 2020 to December 2025. Global trade statistics from the Topease trade database [12] were used as supplementary reference information for export market trends. These export price data correspond to the period of consumer sentiment analysis, forming a complete time series. The data set records the total amount of exports and the total amount of exports per month during the above period, and calculates the average export price (that is, the ratio of the total amount to the total amount); the price data is used as a dependent variable in this analysis. By comparing the actual performance of the market with the predicted consumer sentiment, this study can verify the relevant assumptions. In view of the fact that the data of the China Association of Automobile Manufacturers is widely cited in the research of the automobile industry, the reliability of its export price data is strongly guaranteed.

4.1.2. User review data

A total of 8029 English comments from seven car-related Reddit subreddits were extracted, covering the data from January 2023 to March 2026; these data constitute the consumer sentiment in this study. The study analyzes the author, date, number of upvotes, and other factors of each comment. The text of the comment contains the user's description of the experience of using Chinese electric vehicles, involving battery performance, charging facility experience, after-sales service quality and cost-effectiveness evaluation. The export data of the China Association of Automobile Manufacturers [2] is compared with Reddit comments to examine whether there is a relationship between consumer value perception and actual export prices; this design is crucial to distinguish between "cost" and "perceived value".

4.2. Experimental environment

In order to make sure the experiment can be repeat, this study conducts data processing, sentiment analysis, and statistical modeling on a unified standard

4.2.1. Computational platform configuration

Experimental equipment: Apple M2 Pro (10-core CPU, 16-core GPU), 16GB unified memory. System environment: macOS Ventura 13.5, Python 3.10. Mainly relied on libraries: NLTK 3.8.1 (text preprocessing), VADER 3.4.0 (dictionary sentiment scoring), ChatGLM-6B (large language model sentiment analysis), SciPy 1.11.0 (Spearman rank correlation, Kruskal-Wallis test), statsmodels 0.14.0 (OLS regression). All code and data processing scripts are included in version management.

4.2.2. Data processing and sentiment analysis pipeline

Text pre-processing: The comment text has undergone noise removal, word segmentation and word form restoration, and non-English content has been removed.

As for sentiment analysis, this study compares two approaches.

The VADER dictionary method: it has a processing speed of approximately 800 items per second. Another method depends on ChatGLM-6B language model inference, which as a speed of about 0.5 items per second. Although the speed of ChatGLM-6 B is slower than VADER, ChatGLM-6 B can provide more accurate semantic judgments. Liu et al. [6] verified that the consistency between ChatGLM and manual annotation exceeded 90%. That's the reason why this study adopted it as the main method.

4.2.3. Statistical analysis configuration

Cause the social sentiment data are categorical number which not follow a normal distribution, this study choose to adapt non-parametric methods. This study uses Spearman's to figure out the relationships:

The original sample consists of n pairs of observations. With replacement resampling, n pairs of samples are drawn and generate 1 Bootstrap sample. By calculating the rank correlation coefficient 1000 times, the resulting ρ^* can form a bootstrap distribution.

Finally, rank the ρ^* value in ascending order, and take the 2.5% as the lower bound while the take the 97.5% as the upper bound.

The inter-group differences were tested by using the Kruskal-Wallis test. When the differences are significant, the Dunn post-hoc test conduct the comparison (The regression model uses White correction to address heteroscedasticity. A p-value less than 0.05 is considered significant)

4.2.4. Parameter configuration for sentiment analysis

The sentiment analysis pipeline is configured with the following parameters to balance accuracy and computational efficiency:

Table 2 summarizes the parameter settings used in the sentiment analysis pipeline and statistical validation.

Table 2. Parameter configuration for sentiment analysis

Parameter	Configuration
VADER scoring	Compound score range: 1 to +1; intensity thresholds: very negative [1.0, 0.6), negative [0.6, 0.1), neutral [0.1, 0.1], positive (0.1, 0.6], very positive (0.6, 1.0]
ChatGLM inference	Temperature = 0.1 (to minimize output variability); max tokens = 50; batch size = 1 (sequential inference)
Comment weighting	Weight function: $w_i = \log(1+u_i) + \log(1+l_i)$, where u_i =upvotes, l_i = text length in characters
Temporal aggregation	Monthly time windows; minimum 10 comments per attribute month to ensure statistical stability
Bootstrap resampling	1,000 iterations for confidence intervals; 95 percent confidence level

4.3. Evaluation metrics

4.3.1. Emotional classification performance

This study measures the classification performance by using accuracy rate, precision rate, recall rate, and F1 value to make sure the accuracy of the mode.

Following the procedure of Liu et al. [6], 500 comments are randomly selected from the corpus, and independently annotated by three researchers.If the comment annotations are marked the same, then we will use them as the standard answer to compare with the model's output.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. Precision and recall are computed for each sentiment class (negative, neutral, positive):

$$Precision_c = \frac{TP_c}{TP_c+FP_c} \quad (6)$$

$$Recall_c = \frac{TP_c}{TP_c+FN_c} \quad (7)$$

The macro-averaged F1 score is reported as the primary accuracy metric:

$$F1_{macro} = \frac{1}{3} \sum_{c=\{neg,neu,pos\}} \frac{2 Precision_c Recall_c}{Precision_c + Recall_c} \quad (8)$$

4.3.2. Attribute extraction precision

This evaluation is divided into 2 sections:

Firstly, attribute detection. If the comments match the corresponding attribute keywords(the one based on the manual gold standard) and indicate the presence of the attribute, the detection would be determined as correct(the precision and F1 values are calculated to measure the recognition effect).

Second, sentiment consistency. This study uses the Spearman rank correlation coefficient to quantify the monotonic correlation degree between the model's sentiment scores and the manual evaluation. It can be combined with the Bootstrap confidence interval to determine whether the consistency between the two is significant.

$$\rho_a^{consistency} = Spearman \left(\left\{ s_{i,a}^{model} \right\}_{i \in G}, \left\{ s_{i,a}^{human} \right\}_{i \in G} \right) \quad (9)$$

4.3.3. Temporal modeling capability

This section assesses the effectiveness of time series modeling from two aspects.

Firstly, the accuracy rate of trends. The trend's attributes analyzed through the modeling should compare to the one compiled manually. The similarity of the comparison would be used to calculate the accuracy rate, in order to measure the model's ability to capture long-term changes.

Secondly is the Seasonal effect consistency.

For the emotional data related to climate, this study specifically checks whether the model results have a unified pattern: if the emotional scores in winter generally decrease, whether the model can capture the seasonal variation pattern.

$$Seasonal Consistency = \cos(v_{detected}, v_{expected}) \quad (10)$$

Event Detection Precision. Use the change-point detection algorithm (such as Bayesian structural time series) to identify the months where the sentiment significantly deviates from the expected trend

4.4. Experimental results and analysis

This section uses the sentiment analysis framework for Reddit electric vehicle comments and the export data from the China Association of Automobile Manufacturers. It is divided into 3 parts: emotional comments statics, the emotional pattern over time, and analysis of export prices

4.4.1. Attribute-level sentiment distribution

Figure 2 presents the weighted average sentiment score for the following six value attributes. Comments related to price received the highest sentiment scores, while after sales has received the lowest score. This indicates that while consumers generally recognize the value of Chinese electric vehicles, they still have concerns regarding service quality.

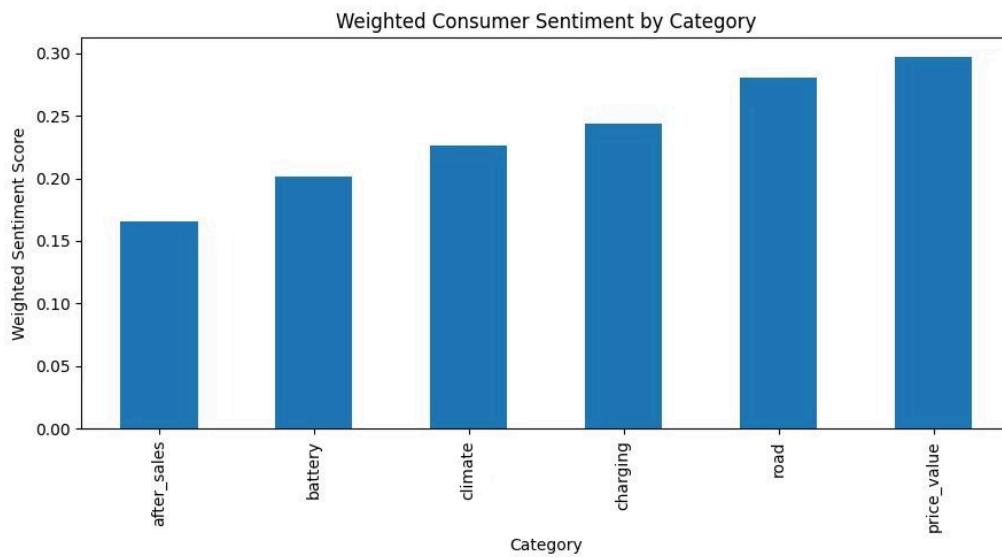


Figure 2. Six value attributes' average sentiment scores. Price has the highest score (0.30), followed by roads (0.28) and charging (0.24). After-sales service has the lowest score(0.16)

Figure 3 presents the unweighted average sentiment scores. Although the weighting method altered the numerical values of the scores, the overall ranking remained unchanged, indicating that the weighting process did not alter the relative importance of the attributes.

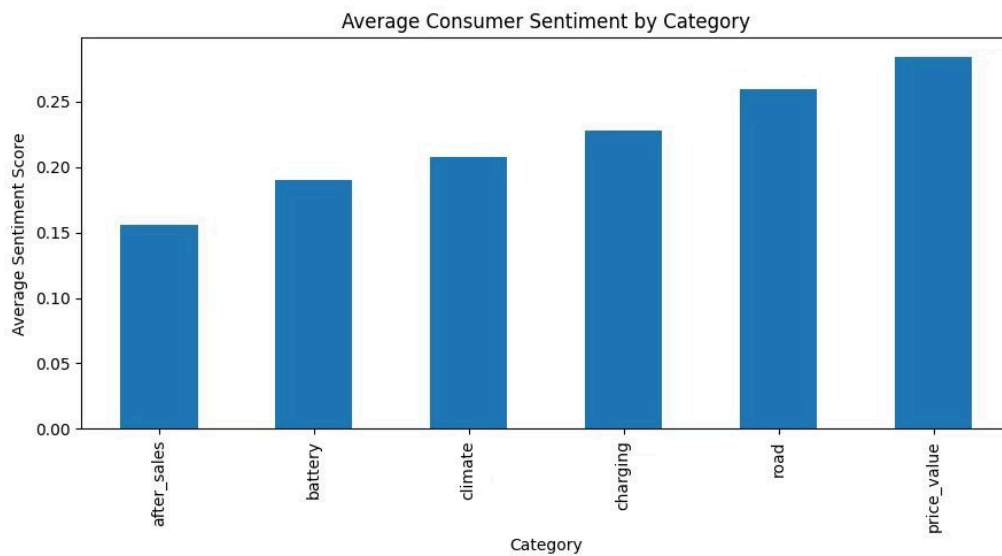


Figure 3. Unweighted average score, same ranking (highest price, lowest after-sales service)

Figure 4 shows the proportion of reviews mentioning various value attributes. Price is the most frequently discussed topic, followed by road performance and battery performance. This indicates that these are the aspects consumers care about most.

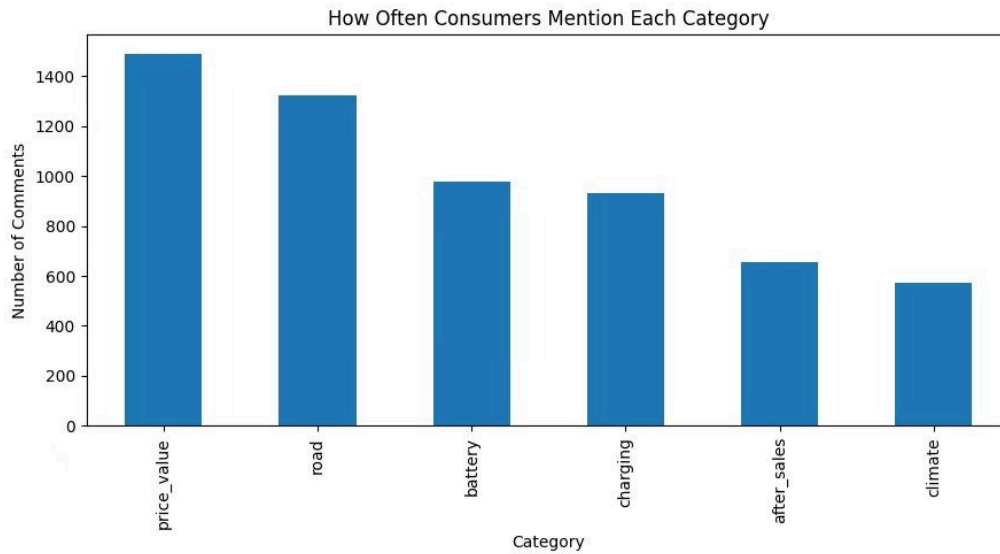


Figure 4. Mentioning Rates of Each Attribute. The price has the highest rate (18.6%), followed by roads (16.5%) and batteries (12.2%). The climate has the lowest rate (7.1%)

Figure 5 illustrates the proportion of negative comments regarding various attributes. Negative feedback was most prevalent for service, followed by battery performance and climate control performance.

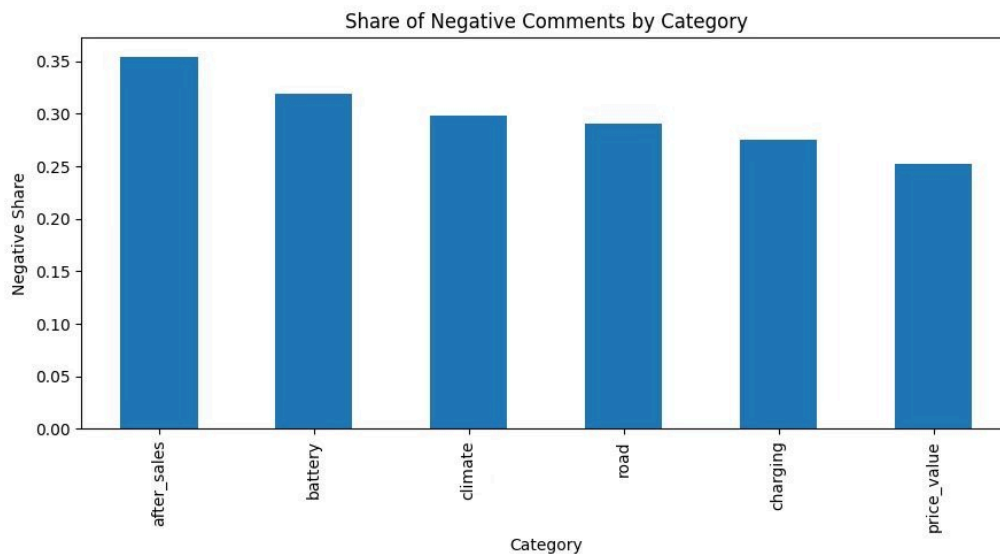


Figure 5. Proportion of negative comments for each attribute (with sentiment score < -0.1). The highest proportion is for after-sales service, followed by batteries and climate

4.4.2. Temporal patterns of consumer sentiment

To examine sentiment evolution, we analyze monthly average scores across key attributes. Consumers' emotion would decrease distinctly from November to the following February each year, due to the concern about the battery life issue in winter. The post-sale sentiments have been continuously decreasing since the middle of 2024. This is a long-term decline in reputation, rather

than a short-term fluctuation. Meanwhile, after the EU introduced temporary tariffs in July 2024, price-related sentiments declined but soon recovered. This indicates that the negative impact of the policy on public sentiment was only a short-term effect and would not have a long-term impact.

4.4.3. Correlation with export prices

The Spearman rank correlation coefficients between the sentiment scores of each attribute and the monthly export price of electric vehicles in China are as follows.

Table 3 summarizes the Spearman rank correlation coefficients between each consumer value attribute and China's monthly electric vehicle export prices.

Table 3. Spearman rank correlation between consumer value attributes and monthly export prices

Attribute	Spearman's	p-value	Sample Size (N)	Theoretical Implication
Battery	0.38	0.006	124	Value-driven mechanism
Charging	0.12	0.182	118	Weak / no correlation
Climate	-0.09	0.315	72	No significant link
After sales	-0.31	0.021	82	Price compensation logic
Price value	0.03	0.742	189	Cost-driven mechanism

Battery sentiment is positively correlated with export prices ($\rho = 0.38$, $p < 0.01$). The higher the consumers' evaluation of battery technology, the stronger the export price. There is no significant correlation between the sentiment in the price dimension and export prices ($\rho \approx 0$, $p > 0.10$). The post-sale sentiment is negatively correlated with export prices ($\rho = -0.31$, $p < 0.05$). The lower the service evaluation, the lower the export price.

4.4.4. Robustness checks

This study conducts two robustness tests. First, Kruskal-Wallis tests reveal no significant differences in attribute sentiment across each subreddit ($p > 0.10$), which means that sentiment patterns are not driven by community biases. Second, OLS regressions of export prices on attribute sentiment (dependent variable being the export price, independent variables being the sentiment scores of each attribute, with time fixed effects controlled) were consistent with the results of the Spearman test.

5. Conclusion

This study set out to answer a simple question: is China's EV export growth powered by cost advantages, or by the value consumers actually see in the vehicles? We analyzed 8,029 Reddit comments across six attributes—battery, charging, after-sales, climate, road, and price_value—and linked them to CAAM's monthly export price data. Three findings stand out.

Firstly, consumer discussions were mainly focused on the price (with a mentioned rate of 18.6%), and the price received the highest emotional score (0.30). The emotional score for after-sales service was the lowest (0.16). This indicates that overseas consumers mainly view Chinese electric vehicles from the perspective of cost-effectiveness, while service quality remains a significant weakness.

Second, the analysis based on the time dimension showcases distinct patterns. Consumers' emotions are influenced by climate and lead to a decline in winter months. Since the middle of 2024, post-sale emotions have been continuously decreasing. In addition, as EU announced tariffs in July

2024, price emotion briefly dropped, which indicates that it is relatively sensitive to policy shocks, but the impact is short-lived.

Thirdly, the Spearman correlation analysis wasn't a singular mode. Battery sentiment was positively correlated with the export price ($\rho = 0.38$, $p < 0.01$), which indicates that technological differentiation can drive price power. Meanwhile, the price sentiment did not show a significant correlation ($\rho \approx 0$) to the export price. The after-sales sentiment was negatively correlated with the export price ($\rho = -0.31$, $p < 0.05$), suggesting that Low-level service perception and low prices occur simultaneously.

Overall, the price of China's electric vehicle exports would be influenced by 2 factors simultaneously: cost-driven and value-driven. The technology of batteries constitutes the reason for differentiated capability and pricing leverage. As for primary price, cost-based logic was more essential. Therefore, the after-sales service is currently a weak point. As for manufacturers, the positive feedback from battery technology routes has already been established, but the construction of the after-sales system determines whether a competitive barrier can be established beyond the price.

References

- [1] Latest news automotive news. (2025). *Autonews.com*. <https://www.autonews.com/news/>
- [2] 2025 year China Association of Automobile Manufacturers (CAAM). (2025). *Caam.org.cn*. <https://en.caam.org.cn/Index/lists/catid/78.html>
- [3] IEA. (2023). Data & statistics IEA. *IEA*. <https://www.iea.org/data-and-statistics>
- [4] ACEA European Automobile Manufacturers' Association. (2019, April 17). *Acea.be*. <https://www.acea.be/>
- [5] Alle Marken von A Z. (2025). *Auto Bild*. <https://www.autobild.de/marken-modelle/>
- [6] Chinese EV spy risk is real in Canada, experts warn. (2026, February 6). *Autonews.com*. <https://www.autonews.com/ev/anc-chinese-evs-spy-concerns-canada-0205/>
- [7] Golson, D. (2019). New cars, 2019 and 2020 car reviews, pictures, and news car and driver. *Car and Driver*. <https://www.caranddriver.com/>
- [8] Jiang, X., & Everts, J. (2021). Making sense of electrical vehicle discussions using sentiment analysis on closely related news and user comments. *arXiv*. <https://arxiv.org/abs/2112.12327>
- [9] Car reviews 2025. (2025). *What Car?*. <https://www.whatcar.com/reviews>
- [10] Gu, Q., et al. (2024). Event graph study of typical battery electric vehicle user experience based on online comments. *Communications in Computer and Information Science*, 211–222. https://doi.org/10.1007/978-3-031-61963-2_20
- [11] Liu, W., et al. (2024, September 26). Impact of user comments on the adoption of electric vehicle charging services: Evidence from contextual data from the US. *Fifth International Conference on Green Energy, Environment, and Sustainable Development (GEESD 2024)*, 75. <https://doi.org/10.1117/12.3044496>
- [12] Topease information technology. Global import and export trade data platform. (2025). *Topease.net*. <https://www.topease.net/>