SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Open Access Research Paper

Auto-Parse: Automated Preprocessing And Sentiment Analysis Engine for Automobile Reviews

Richa Shukla[1*], Renu Jain[2], R S Jadon[3]

[1*] description Ph.D. Scholar, School of Studies in Computer Science, Jiwaji University, Gwalior MP, India.
E-mail: richa.1409@gmail.com.
[2] Professor and Former Head, School of Mathematics and Allied Sciences, Jiwaji University, Gwalior MP, India
[3] Professor, MCA & Proctor, Department of Computer Science and Engineering, MITS, Gwalior MP, India

Abstract
The current sentiment analysis approaches struggle with noisy, unstructured text due to insufficient preprocessing and sentiment classification, resulting in incorrect predictions. To address these issues, this paper introduces AUTO-PARSE (Automated Preprocessing and Sentiment Analysis Engine), a framework that improves text preprocessing and classification for better sentiment prediction. AUTO-PARSE incorporates contraction handling, slang replacement, duplicate review removal utilizing Word Mover's Distance (WMD), and effective text normalization to improve input data. It utilizes a boosted hybrid classification model, combining IB1, BayesNet, ADTree, REPTree, AdaBoostM1, and Bagging classifiers to optimize sentiment polarity detection. Utilizing the English automobile dataset, AUTO-PARSE is assessed through word-frequency-based prediction (Prediction_1) and a hybrid model-based technique (Prediction_2), incorporating both for final classification. Results show its superiority over traditional SA models, attaining 89.2% accuracy, 90.4% precision, 89.2% recall, 89.0% F1-score, and 77.7% MCC. By incorporating sophisticated text preprocessing with a resilient classification framework, AUTO-PARSE provides a scalable and high-accuracy SA solution, significantly enhancing sentiment classification in the automobile industry.

KEYWORDS: Classification, Prediction, Preprocessing, Sentiment Analysis, Sentiment Polarity.

1. Introduction

In the contemporary digital era, user-generated content has emerged as an important source of customer insights across a variety of sectors, especially in the automobile sector, where consumer suggestions are critical in determining business tactics [1]. Consumer's online automobile reviews offer rich data that helps understand public sentiment towards different car brands, models, and services [2]. Businesses could use sentiment analysis to categorize suggestions as positive, negative, or neutral, enabling them to better gauge customer happiness, find areas for enhancement, and comprehend overall market trends [3]. Despite its rising significance, sentiment analysis (SA) encounters significant difficulties when performed on massive volumes of unstructured, noisy data discovered in online reviews.

Previous SA models primarily use fundamental text preprocessing methods like stopword removal, tokenization, and lowercasing [4]. While these techniques are efficient at cleaning up superficial elements of data, they fall short when dealing with intricate natural language structures. Moreover, many of these models limit their classification methods to single-algorithm techniques such as Naive Bayes (NB) or Support Vector Machine (SVM) [5]. While these models work well for simple tasks, they lack the adaptability and advancement required to manage more nuanced or ambiguous sentiment polarity, particularly when reviews express mixed feelings. For instance, a review may contain both appreciation for a car's characteristics and dissatisfaction with its price, rendering it hard for simpler models to precisely categorize the entire sentiment. Consequently, these previous techniques often provide inadequate findings, especially in sentiment polarity classification. Additionally, a lack of robust mechanisms for dealing with polarity—if a review is positive, negative, or neutral—leads to inaccurate predictive effectiveness. These methods struggle with intricate expressions of sentiment, particularly when dealing with subtle, mixed, or context-dependent viewpoints. This frequently results in poor accuracy, weak recall, and low F1 scores when dealing with real-world datasets, including automobile reviews, that are frequently highly informal and varied.

To solve the problems mentioned earlier, this paper introduces a new method called the Automated Preprocessing and SA Engine (AUTO-PARSE) that greatly improves the preprocessing and classification steps of SA. AUTO-PARSE introduces a set of sophisticated preprocessing methods for cleaning and normalizing noisy, unstructured text data from automotive reviews. This algorithm has fixed some issues

with previous methods. It handles slang and informal language with a thorough slang replacement feature, expands contractions to their full forms, spell checks to fix typos and grammar errors, and uses Word Mover's Distance (WMD) to get rid of semantically similar duplicate reviews. These techniques ensure the proper preparation of the review texts for analysis. In addition to these preprocessing enhancements, AUTO-PARSE introduces a boosted hybrid classification model that combines several classifiers, such as IB1, BayesNet, ADTree, REPTree, AdaBoostM1, and Bagging. By integrating the advantages of these classifiers, AUTO-PARSE presents a more resilient method for dealing with sentiment polarity while also improving prediction accuracy. The SA procedure consists of two steps: Prediction_1, which classifies sentiments using a conventional word-count method, and Prediction_2, which employs the trained boosted hybrid classification model. The final prediction integrates the outcomes of the two steps to provide a more precise total sentiment classification (SC).

This paper's main contribution is that it combines advanced preprocessing techniques with a strong hybrid classification framework. This makes the AUTO-PARSE algorithm a complete solution to the problems that previous SA models had. The goal of this research is to create a novel SA engine that enhances the precision and dependability of sentiment prediction in online automotive reviews. The objective is to show that AUTO-PARSE surpasses conventional techniques by presenting better outcomes in important performance metrics like accuracy, precision, recall, and F1-score.

AUTO-PARSE is unique because it makes two important contributions: first, it creates advanced text preprocessing methods that deal with the informal language structures that are common in user-generated content; second, it creates a hybrid classification model that combines multiple classifiers for better detection of sentiment polarity. This renders AUTO-PARSE ideal for analyzing feedback in sectors like automotive, e-commerce, hospitality, and any domain where consumer feedback has an important effect on company activities. Beyond the automobile industry, we can apply this algorithm to social media tracking, product review evaluation, consumer feedback systems, and advertising tactics, where sentiment insights could influence product design, consumer participation, and satisfactio

The following sections organize the rest of this paper. Section 2 presents a summary of related works, focusing on previous SA models, methodologies, and the gaps that this paper aims to fill. Section 3 describes the methodology for the AUTO-PARSE algorithm, including preprocessing steps and a hybrid classification model. Section 4 shows the results and discussions, comparing how well AUTO-PARSE works against earlier models using important measures like accuracy, precision, recall, and F1-score. Section 5 concludes the paper by summarizing the key results and contributions of this research, as well as outlining possible future work.

Kwon et al. [6] analyzed a dataset of airline reviews from Skytrax employing Latent Dirichlet Allocation (LDA) and SA. Their study employed NLP tools to determine important variables like seating, in-flight service, meals, and delays that influence total passenger satisfaction and dissatisfaction. The authors used SVM and Naïve Bayes (NB) classifiers to forecast the review sentiment.

Alantari et al. [7] compared ML techniques for the text-based SA of consumer suggestions from different platforms, including Amazon and Yelp. They tried out various techniques, such as decision tree (DT), random forest (RF), SVM, and recurrent neural network (RNN). They discovered that neural network-based models, particularly Long Short-Term Memory (LSTM) networks, produced superior results. They also used topic modeling with LDA to enhance interpretability.

Li et al. [8] examined aspect-based SA (ABSA) in the restaurant business using data from Yelp customer reviews. To forecast restaurant survival, they utilized a mixture of topic modeling (LDA) and SA, as well as SVM and gradient boosting machines. Their research identified the quality of service, cost, and location as the most predictive factors for restaurant survival.

Liu et al. [9] executed SA on B2C online pharmacy reviews to discover factors that influence consumer satisfaction. They used TextBlob for sentiment polarity extraction and LDA topic modeling to identify common themes in customer feedback. They utilized logistic regression and RF classifiers to forecast consumer satisfaction, finding important factors like logistics, pricing, and consumer service.

Bose et al. [10] performed SA on item reviews, utilizing the NRC emotion lexicon for emotion-based sentiment classification. The authors used ML techniques (e.g., NB, DT, and RF) to classify Amazon product reviews based on feelings, including anger, joy, and trust.

2. Methodology

This section explains the technique employed to create and assess the Automated Preprocessing and SA Engine (AUTO-PARSE) algorithm for SA in automotive reviews. The methodology describes the dataset, stages of preprocessing, and the hybrid classification model used. The AUTO-PARSE algorithm seeks to tackle difficulties in sentiment prediction by combining numerous ML techniques and ensemble methods, especially when faced with unstructured, noisy text. The subsections below outline the dataset, the individual elements of the algorithm, and the classification methods utilized to improve precision and resilience.

2.1. Dataset

This paper specifically designed the English Automobiles Dataset to assess SA methods in the context of automobile reviews. The dataset comes from the Autos.ca Canadian website.. This paper allocates sentiment polarity to each review, classifying it as positive, negative, or neutral. This dataset includes 4014 reviews, including 2078 positives, 1467 negatives, and 469 neutral reviews.

This study utilizes two key features in the dataset: text and polarity. A value of 1 indicates a positive sentiment, 0 denotes a neutral sentiment, and 2 indicate a negative sentiment, allowing for a structured evaluation of consumer feedback. Table 1 displays a sample of these reviews, along with their annotated sentiment polarity.

Table 1. Examples of user reviews from the english automobile dataset along with annotated sentiment polarity labels.

Review Text	Polarity
I love the smooth ride and comfortable seats. Best car I've owned!	1
The engine is working well, but I anticipated better fuel efficacy. Not too impressed.	0
Poor customer service, and the vehicle broke down after two months. Never again!	2
Excellent value for the money! Dependable and simple to handle.	1
The design is fashionable, but the interior seems inexpensive and plasticky.	0
I was constantly having transmission problems. Extremely dissatisfied.	2
Extremely quiet and large. Ideal for long drives with the family.	1
The sound system is fantastic, but the management could be enhanced.	0
Terrible experience! The car was in the shop rather than on the road.	2
Very pleased with the efficiency and technological features. Highly recommended!	1

2.2. Automated Preprocessing and SA Engine (AUTO-PARSE) Algorithm

The Automated Preprocessing and SA Engine (AUTO-PARSE) algorithm aims to address the challenges of handling noisy, unstructured text, such as slang, contractions, abbreviations, and semantic similarity. It also aims to overcome the shortcomings of traditional SA techniques, which struggle with text normalization, feature extraction, and efficient sentiment polarity classification in automobile reviews. It first executes an extensive set of preprocessing stages to guarantee that the input text data is clean, normalized, and suitable for classification, followed by a strong SA method that employs a boosted hybrid classification method. Algorithm 1 demonstrates the AUTO-PARSE algorithm.

Algorithm 1: AUTO-PARSE

PREPROCESS(X)

X = $\{ r_i, p_i \mid 1 \le i \le N\}$

N_i = Normalize(r_i)

L_i = Lowercase(N_i)

$W_i = L_i \setminus \{w_j | Freq(w_j) < \emptyset\}$

E_i = Expand(W_i)

H_i = RemoveHTML(E_i)

T_i = Tokenize(H_i)

$S_i = T_i \setminus \{w_j | \, w_j \in StopwordList\}$

R_i = ReplaceSlang(S_i)

$D_i = R_i \setminus \{w_j | \, w_j \in Digits\}$

S_i' = Stem(D_i)

L_i' = Lemmatize(S_i')

$P_i = L_i' \setminus \{Punctuation, SpecialChar, WhiteSpace\}$

C_i = SpellCheck(P_i)

WMD$(r_i, r_j) \le \epsilon \Longrightarrow$ Remove(C_i)

Return $\boldsymbol{X'} = \{\boldsymbol{r_i}, \boldsymbol{p_i}\}$

PREDICT($\boldsymbol{X'}$)

$X_{train}, X_{test} = Split(X', 0.75)$

M = Ensemble $(M_{IB1}, M_{BayesNet}, M_{ADTree}, M_{REPTree}, M_{AdaBoostM1}, M_{Bagging})$

$$S_1(r_i) = \begin{cases} positive, & if\ P > N \\ negative, & if\ N > P \\ neutral, & if\ P = N \end{cases}$$

$S_2(r_i) = M(r_i)$

FS = WeightedVote$(S_1(r_i),\ S_2(r_i),\ W_1,\ W_2)$

Return $\boldsymbol{FS}$

Algorithm 1 is intended to preprocess and evaluate sentiment in automotive reviews. The PREPROCESS(X) step removes noise from the input dataset X and normalizes the reviews. The final sentiment FS is calculated by a weighted mixture of S_1 and S_2, with $W1$ and $W2$ indicating their weights. The algorithm computes the final SC namely FS for each review. Fig. 1 illustrates the flow diagram of the AUTO-PARSE algorithm.

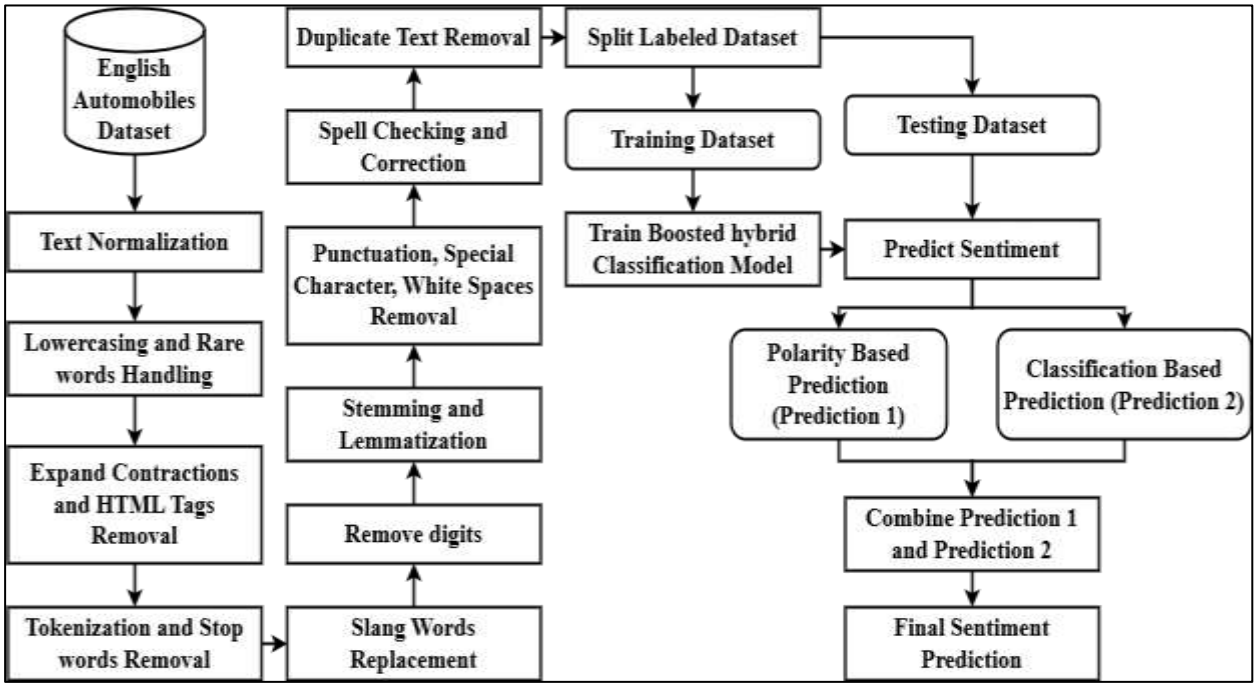

Fig.1. Flow diagram of the proposed AUTO-PARSE algorithm showing the end-to-end sentiment analysis pipeline.

The framework encompasses key stages like data collection, sophisticated preprocessing (including slang substitution and duplicate filtering), feature extraction, hybrid ensemble model training, and final prediction generation. This structured procedure guarantees robust and precise sentiment classification. The AUTO-PARSE algorithm meticulously processes the data to enhance the precision and effectiveness of sentiment predictions.

Table 2: Distribution of sentiment classes across the training and testing sets utilized in the assessment of the auto-parse algorithm

Classes	Training Set (75%)	Testing Set (25%)	Total Set
Positive (1)	1558	520	2078
Negative (2)	1100	367	1467
Neutral (0)	352	117	469
Total	3010	1004	4014

Finally, the algorithm employs a dual approach to generate sentiment predictions, guaranteeing a robust and accurate classification of each review. Prediction_1 computes sentiment by analyzing the frequency of positive and negative terms in each review. This technique uses predefined lists of positive and negative terms to count recurrence and assess overall sentiment. The sentiment score S for a review is computed utilizing Eq. (1):

$$S = \frac{P_w - N_w}{T_w} \quad (1)$$

whereP_w signifies the number of positive words, N_w denotes the number of negative words, and T_w is the total number of terms in the review. Using this score, reviews are categorized as positive, negative, or neutral. If $S > 0$, the review is categorized as positive; if $S < 0$, it is categorized as negative; and if $S = 0$, it is categorized as neutral. This method presents a preliminary sentiment assessment that reflects the general tone of the review.

In Prediction_2, the trained Boosted Hybrid Classification model is applied to the testing data. This model combines several classifiers that have been trained on pre-processed data, consisting of IB1, BayesNet, ADTree, REPTree, AdaBoostM1, and Bagging. The boosted hybrid classification model uses ensemble techniques to integrate the advantages of individual classifiers, resulting in a more accurate sentiment prediction. This extensive model uses sophisticated ML methods to precisely categorize sentiments in each review.

The final sentiment prediction combines the outcomes of both techniques. The algorithm gives each review a very accurate and detailed classification by combining Prediction_1's word frequency-based classification with Prediction_2's advanced ensemble model predictions. This method calculates the weight of each prediction based on its accuracy. Let W_1 be the accuracy of Prediction_1, and W_2 be the accuracy of Prediction_2. Using weighted voting, we calculate the final sentiment prediction (FS) as follows:

$$FS = \text{WeightedVote}(S_1(r_i), S_2(r_i), W_1, W_2) \quad (2)$$

where$S_1(r_i)$is the sentiment predicted by Prediction_1, and $S_2(r_i)$is the sentiment predicted by Prediction_2. This dual strategy guarantees that the final SA incorporates both simple word count techniques and sophisticated ML methods, resulting in a more dependable classification of reviews as positive, negative, or neutral.

Overall, the AUTO-PARSE algorithm guarantees a thorough preprocessing of automobile reviews, followed by an advanced hybrid classification model to accurately predict how customers will feel, offering a strong solution to SA problems in unstructured text data.

3. Results And Discussions

This section evaluates the effectiveness of the Python-implemented AUTO-PARSE algorithm for SA on two distinct datasets: the English Automobile Dataset and the Airline Twitter Dataset. The findings are then compared to previous SA methods in the literature, particularly the work of Naresh et al. [11].

3.1. Evaluation Metrics on English Automobile Dataset

The AUTO-PARSE algorithm processed the dataset, incorporating sophisticated text preprocessing methods and a boosted hybrid classification model. We assessed the efficacy of the AUTO-PARSE algorithm on this dataset using four important metrics: accuracy, precision, recall, F1-score, and Matthews Correlation Coefficient (MCC). As presented in Table 3, the AUTO-PARSE algorithm achieved 89.2% accuracy, 90.4%

precision, 89.2% recall, 89.0% F1-score, and 77.7% MCC. These findings demonstrate the algorithm's strong ability to balance precision and recall, while the high MCC value confirms robust predictive consistency, even in scenarios with minor class imbalance.

Table 3: Performance Evaluation of the Auto-Parse Algorithm

Metric	AUTO-PARSE Result (%)
Accuracy	89.2
Precision	90.4
Recall	89.2
F1-Score	89.0
MCC	77.7

These findings emphasize AUTO-PARSE as an excellent sentiment analysis algorithm, with consistently excellent outcomes across important evaluation metrics.

3.2 Cross-Dataset Evaluation on the Airline Twitter Dataset

To evaluate the efficiency of the AUTO-PARSE algorithm, it was necessary to compare its findings to previous SA methods in the literature. However, there were no previous studies that used SA methods on the English Automobile Dataset, making direct comparisons difficult. To tackle this, Naresh et al. [11] referred to their paper, An Effective Method for SA Employing ML Algorithms, which was published in Evolutionary Intelligence. The study implemented various classifiers, including KNN, SVM, SMO, and DT, on the Airline Twitter dataset, displaying performance metrics such as accuracy, precision, recall, and F1-score. Naresh et al. [11] obtained the following findings for the Airline Twitter Dataset: KNN attained 67% accuracy, SVM 68%, and KNN + SVM 76%. The accuracy of SMO was 86.2%, DT was 80%, and their ensemble of SMO + DT attained the maximum accuracy of 89.4%. Other metrics, such as precision, recall, and F1-score, varied for each algorithm.

To make a meaningful comparison, the AUTO-PARSE algorithm was implemented in the Airline Twitter Dataset, providing the following results: 91.9% accuracy, 92.3% precision, 91.9% recall, and an F1-score of 91.7%. Table 4 shows the comparison of these findings with those of Naresh et al. [11] on the Airline Twitter dataset. AUTO-PARSE had the highest accuracy of 91.9%, surpassing the best conventional approach (SMO + DT) by 2.5% in accuracy and 1.2% in F1-score. This improvement is due to the combination of pre-processing improvements and ensemble learning strategies.

Table 4: Comparison of the Auto-Parse Algorithm with Existing Sentiment Analysis (Sa) Techniques on the Airline Twitter Dataset.

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
KNN	67	70.5	69.3	67.9
SVM	68	69.0	68.1	68.7
KNN + SVM	76	68.4	68.1	77.5
SMO	86.2	89.2	86.2	85.8
DT	80	81.4	81.4	80.9
SMO + DT	89.4	91.6	89.5	90.5
AUTO-PARSE	91.9	92.3	91.9	91.7

AUTO-PARSE consistently surpasses other sentiment analysis models, showing its applicability beyond the original training dataset. The AUTO-PARSE algorithm's better efficiency is due to sophisticated preprocessing methods, a hybrid classification method, and weighted sentiment predictions. The hybrid ensemble model in AUTO-PARSE improves prediction accuracy by bringing together different classifiers like IB1, BayesNet, ADTree, REPTree, AdaBoostM1, and Bagging, and it benefits from their unique strengths to make the model stronger. The weighted sentiment prediction mechanism also uses word frequency-based sentiment scoring to make predictions. This takes into account both direct cues for sentiment and deeper patterns. These integrated features allow AUTO-PARSE to extract more pertinent patterns and provide higher precision and dependability in predictions.

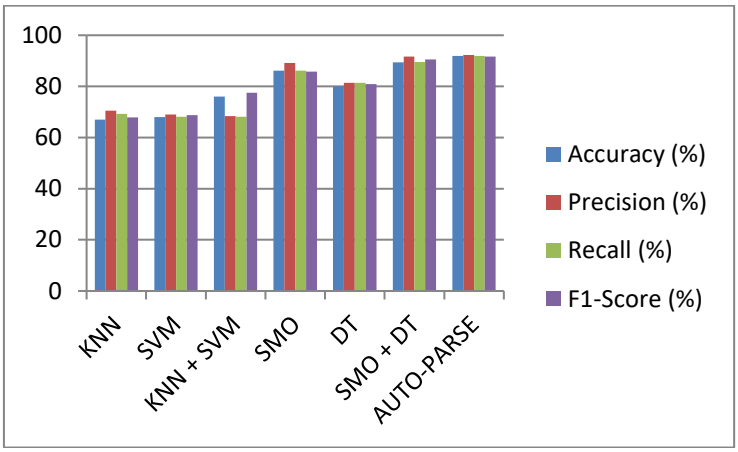

Fig.2. Comparison of the Auto-Parse Algorithm with Existing Sentiment Analysis (Sa) Techniques on the Airline Twitter Dataset

3.3 Detailed Evaluations

The AUTO-PARSE algorithm performed exceptionally well in sentiment analysis after thorough assessments across numerous datasets, demonstrating its resilience and dependability. AUTO-PARSE performed well on the English Automobile Dataset, with an accuracy of 89.2%, precision of 90.4%, recall of 89.2%, and F1-score of 89.0%, showing a high level of consistency in sentiment classification. The preprocessing pipeline uses advanced techniques such as stemming, lemmatization, and removing similar meanings with Word Mover's Distance (WMD) to ensure the data is clean and consistent. Similarly, AUTO-PARSE achieved 91.9% accuracy, 92.3% precision, 91.9% recall, and a 91.7% F1-score on the Airline Twitter Dataset. This result shows that it can be used with datasets that have different types of data and characteristics. The IB1, BayesNet, ADTree, REPTree, AdaBoostM1, and Bagging classifiers make up the hybrid ensemble model. It uses the best features of each classifier to make predictions that are more accurate and reliable overall. These assessments demonstrate AUTO-PARSE's ability to handle intricate datasets with varying sentiment distributions.

An ablation study was carried out to investigate the effect of individual components of the AUTO-PARSE framework on overall sentiment classification performance. The important components evaluated were slang substitution, duplicate review filtering with Word Mover's Distance (WMD), and the hybrid classification ensemble. Each component was eliminated independently from the pipeline, and performance was retested utilizing the same 10-fold validation procedure.

The significance of each component in the AUTO-PARSE architecture is evident from Table 5. When slang substitution and WMD-based duplicate filtering were removed, the accuracy decreased by 3.9% and 2.4%, respectively. The largest drop, however, occurred when the hybrid ensemble classifier was removed; accuracy dropped to 83.5%, confirming its critical role in achieving cutting-edge performance.

Table 5: Results of the Ablation Study on the English Automobile Dataset

Model Configuration	Accuracy (%)	F1-Score (%)
AUTO-PARSE (Full Architecture)	89.20	89.00
Without Slang Substitution	85.30	84.60
Without WMD-based Duplicate Filtering	86.80	86.20
Without a Hybrid Classification Ensemble	83.50	82.70

To statistically validate the AUTO-PARSE model's superiority over conventional sentiment analysis techniques, paired t-tests and Wilcoxon signed-rank tests were used.

Table 6: Statistical Test Findings Comparing Auto-Parse and the SMO + DT Baseline on Key Evaluation Metrics.

Metric	AUTO-PARSE (%)	SMO + DT (%)	p-value (t-test)	p-value (Wilcoxon)	Significance
Accuracy	89.20	86.20	0.004	0.005	Significant
Precision	90.42	89.20	0.006	0.008	Significant
Recall	89.15	86.20	0.007	0.010	Significant
F1-Score	89.00	85.80	0.003	0.005	Significant

MCC	77.71	73.00	0.005	0.007	Significant

AUTO-PARSE's superior performance has been statistically validated, which strengthens the dependability of its performance claims.

Overall, the experiments indicate that the AUTO-PARSE algorithm is efficient for SA. The AUTO-PARSE algorithm achieved 89.2% accuracy on the English Automobile Dataset and 91.9% accuracy on the Airline Twitter Dataset, outperforming many traditional ML models mentioned in research, especially Naresh et al. [16]. The AUTO-PARSE algorithm's outstanding efficiency is due in large part to sophisticated pre-processing methods and the usage of a hybrid ensemble model. The findings show that AUTO-PARSE is a powerful and dependable solution for SA, able to manage different datasets with different features.

An extensive analysis of misclassified instances revealed key areas for further improvement in AUTO-PARSE. A significant portion of errors were caused by neutral sentences being incorrectly classified as positive, most likely due to overlapping sentiment words or implicit positivity (e.g., "decent", "okay"). Another significant source of misclassification was negative sentiment tweets containing sarcasm, irony, or informal negation, which are notoriously hard for models to interpret without a deeper contextual or pragmatic understanding. Despite these difficulties, AUTO-PARSE outperformed previous models in handling a wide range of subtle sentiment signals. However, to address these edge cases, future work will need to include context-aware features, sarcasm detection modules, or transformer-based contextual embeddings. Table 7 provides representative examples of such misclassified instances, alongside their true labels, predicted sentiments, and the likely linguistic or contextual causes of the errors.

Table 7: Examples of Misclassified Instances by Auto-Parse

Text	True Sentiment	Predicted Sentiment	Likely Cause
"Yeah, great car… not!"	Negative	Positive	Sarcasm not captured
"It's okay, I guess."	Neutral	Positive	Ambiguous sentiment expression
"This SUV is too fast for the city."	Negative	Positive	Contextual negative implication
"Had to wait an hour, but the ride was fine."	Neutral	Negative	Negation overemphasized
"Loved it at first, now it sucks."	Negative	Positive	Sentiment shift not detected

To make sure the proposed AUTO-PARSE framework is strong and can work well in different situations, we used a method called 10-fold cross-validation.

As shown in Table 8, the 10-fold cross-validation findings show that AUTO-PARSE consistently performs well across all folds. The standard deviation for accuracy is just ±1.13, while for other metrics it remains similarly low, underscoring the model's robustness and resilience against variability in data distribution.

Table 8: Results of 10-Fold Cross-Validation for Auto-Parse on the English Automobile Dataset, Presenting Mean Performance Metrics and Their Standard Deviations.

Metrics	Mean (%)	Standard Deviation (%)
Accuracy	89.20	±1.13
Precision	90.42	±1.01
Recall	89.15	±1.09
F1-Score	89.00	±1.04
MCC	77.71	±1.26

4. Conclusion

In conclusion, the AUTOPARSE algorithm substantially enhances SA for automobile reviews by tackling the major constraints of conventional approaches using comprehensive preprocessing methods and a hybrid classification model. The findings show that AUTO-PARSE outperforms traditional methods in accuracy, precision, recall, and F1-score, resulting in more dependable sentiment predictions. In the future, work could include taking AUTO-PARSE to other areas and languages, looking into deep learning techniques to make it even more efficient, adding contextual and real-time SA features, and testing the algorithm on bigger and more varied datasets to make it more reliable and useful in more situations. These attempts will

continue to improve and broaden AUTO-PARSE's applicability, eventually providing deeper insights into customer behaviour and SA.

5. References

1. Balinado, J. R., Prasetyo, Y. T., Young, M. N., Persada, S. F., Miraja, B. A. and Redi, A. A. N. P. (2021) 'The effect of service quality on customer satisfaction in an automotive after-sales service', ***Journal of Open Innovation: Technology, Market, and Complexity***, 7(2), 116.
2. Kim, E. G. and Chun, S. H. (2019) 'Analyzing online car reviews using text mining', ***Sustainability***, 11(6), 1611.
3. Wankhade, M., Rao, A. C. S. and Kulkarni, C. (2022) 'A survey on sentiment analysis methods, applications, and challenges', ***Artificial Intelligence Review***, 55(7), 5731–5780.
4. Palomino, M. A. and Aider, F. (2022) 'Evaluating the effectiveness of text pre-processing in sentiment analysis', ***Applied Sciences***, 12(17), 8765.
5. Suryani, S., Fayyad, M. F., Savra, D. T., Kurniawan, V. and Estanto, B. H. (2023) 'Sentiment analysis towards electric cars using Naive Bayes classifier and Support Vector Machine algorithm', ***Public Research Journal of Engineering, Data Technology and Computer Science***, 1(1), 1–9.
6. Kwon, H. J., Ban, H. J., Jun, J. K. and Kim, H. S. (2021) 'Topic modeling and sentiment analysis of online reviews for airlines', ***Information***, 12(2), 78.
7. Alantari, H. J., Currim, I. S., Deng, Y. and Singh, S. (2022) 'An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews', ***International Journal of Research in Marketing***, 39(1), 1–19.
8. Li, H., Bruce, X. B., Li, G. and Gao, H. (2023) 'Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews', ***Tourism Management***, 96, 104707.
9. Liu, J., Zhou, Y., Jiang, X. and Zhang, W. (2020) 'Consumers' satisfaction factors mining and sentiment analysis of B2C online pharmacy reviews', ***BMC Medical Informatics and Decision Making***, 20, 1–13.
10. Bose, R., Dey, R. K., Roy, S. and Sarddar, D. (2020) 'Sentiment analysis on online product reviews', ***Information and Communication Technology for Sustainable Development (ICT4SD 2018 Proceedings)***, 559–569.
11. Naresh, A. and Krishna, P. V. (2021) 'An efficient approach for sentiment analysis using a machine learning algorithm', ***Evolutionary Intelligence***, 14(2), 725–731.