electronics

MDPI

Article

Integrating Technical Analysis into Sentiment Analysis: An ASTE Framework for Electric Car Purchase Decision Support Based on LLMs and Semantic BNF

Álvaro Carrasco-Aguilar [1,2,*], M. Mercedes Carmona-Martínez [3], María C. Parra-Meroño [3] and Mar Souto-Romero [4]

1 Facultad de Economía y Empresa, UCAM Universidad Católica de Murcia, 30107 Murcia, Spain
2 Instituto Universitario de Estadística y Ciencia de los Datos, Universidad Complutense, 28040 Madrid, Spain
3 Departamento de Ciencias Sociales, Jurídicas y de la Empresa, Facultad de Economía y Empresa, UCAM Universidad Católica de Murcia, 30107 Murcia, Spain; mcarmona@ucam.edu (M.M.C.-M.); mcparra@ucam.edu (M.C.P.-M.)
4 Departamento de Economía de la Empresa, Facultad de Ciencias de la Economía y Empresa, Universidad Rey Juan Carlos, 28032 Madrid, Spain; mar.souto@urjc.es
* Correspondence: acarrasco73@alu.ucam.edu or acarra10@ucm.es

Abstract: The increasing complexity of purchasing an electric car, influenced by technical specifications and expert reviews, requires advanced Natural Language Processing techniques to extract meaningful insights. This study enhances Aspect Sentiment Triplet Extraction (ASTE) by integrating Large Language Models (LLMs) to identify key aspects, opinions, and sentiments in expert reviews, including technical data traditionally classified as neutral, such as horsepower and battery range. A semantic extension of Backus–Naur Form (BNF) structures input queries for syntactic and semantic accuracy, while a 2-tuple fuzzy linguistic model refines sentiment representation, ensuring interpretability. The proposed model addresses limitations in existing ASTE techniques by incorporating formal grammar structures and linguistic modeling, eliminating the need for complex preprocessing. Applied to expert YouTube reviews of electric cars, the method leverages Google's Gemini model via Python and the Gemini API to rank the top-selling electric cars in the United States. The results confirm the model's effectiveness in aligning technical data with sentiment analysis, making it accessible to non-specialists in Natural Language Processing. This framework enhances decision support in electric car purchases by providing a structured, interpretable, and contextually rich sentiment analysis approach.

Keywords: aspect sentiment triplet extraction; large language models; semantic Backus–Naur Form; fuzzy linguistic models; electric vehicles; expert reviews

check for updates

Academic Editors: Dimitris Apostolou, Francisco Mata and Ramón Carrasco

Received: 7 February 2025
Revised: 25 February 2025
Accepted: 27 February 2025
Published: 4 March 2025

Citation: Carrasco-Aguilar, Á.; Carmona-Martínez, M.M.; Parra-Meroño, M.C.; Souto-Romero, M. Integrating Technical Analysis into Sentiment Analysis: An ASTE Framework for Electric Car Purchase Decision Support Based on LLMs and Semantic BNF. *Electronics* **2025**, *14*, 1020. https://doi.org/10.3390/electronics14051020

1. Introduction

The decision to purchase an electric vehicle is a more complex process than acquiring a conventional car, as it involves additional factors beyond the traditional selection criteria. According to Haustein and Jensen (2018) [1], electric vehicle consumers consider not only rational aspects such as price, range, and access to charging infrastructure, but are also influenced by normative and psychological factors, such as social pressure and environmental norms. Unlike buyers of internal combustion vehicles, who prioritize comfort and familiarity with the technology, potential electric car buyers are often motivated by sustainability and emission reduction. They are also more willing to change their mobility habits based on ecological and social incentives. Since purchasing an electric vehicle involves multiple criteria, the selection process should be approached through a structured framework based on multi-criteria analysis, as highlighted by Więckowski et al. (2023) [2]. This study emphasizes that evaluating an electric car not only requires comparing technical specifications and costs, but also considering environmental and technological criteria to achieve an optimal selection. Expert opinions play a crucial role in this decision-making process, and nowadays, these evaluations are increasingly found on digital platforms and online sources. According to Wijaya et al. (2022) [3], consumers rely on digital advertising and content from platforms such as YouTube to gather information about car models, technical features, and user experiences before making a purchase. These reviews help buyers reduce uncertainty, compare vehicles more efficiently, and better understand how they align with their specific needs.

At this juncture, it is common for buyers to follow a series of systematic steps in this buying process, such as the one proposed in Figure 1 ([1,4]), which presents the structured steps of the purchasing process, detailing the logical progression a buyer follows. It begins with problem recognition, where the need for a product or service is identified, prompting the decision-making process. This is followed by information search, where consumers gather relevant data from various sources, such as expert reviews, online platforms, and product specifications. In the evaluation of alternatives phase, different options are compared using key criteria, considering factors like performance, price, and features. Once an optimal choice is determined, the purchase decision is made, finalizing the transaction. Finally, post-purchase behavior involves assessing satisfaction with the choice, providing feedback, and influencing future purchasing decisions.

Figure 1. Buying process.

This process has already been used in the electric car selection problem [1] using YouTube as source of information. Indeed, according to [5], YouTube is the most influential digital channel during the information search stage of the car purchasing process. Approximately 70% of buyers rely on videos from this platform to gather detailed information about models, prices, features, performance, etc., allowing them to form clear preferences before visiting a dealership. The authors also highlight the importance of expert advice on YouTube, as it plays a key role in guiding consumers' choices during this stage. Pinto [6] highlights YouTube's impact on electric vehicle perceptions by addressing consumer concerns and showcasing benefits. Video blog comments reveal concerns about range, charging, battery life, and costs, while positive sentiments focus on environmental benefits, cost efficiency, and government incentives.

There is also a large textual repository of YouTube electric car reviews at [7] for the purpose of extracting car features through the process known as Aspect Sentiment Triplet Extraction (ASTE), which focuses on identifying specific aspects mentioned in the texts, the opinions related to those aspects, and the sentiment expressed toward them, providing a detailed and structured analysis of the reviews, which are potentially very useful for the evaluation of alternative phases of the aforementioned purchasing process. This study explores key *aspects* such as "performance", "interior-features", and "comfort", among others, and categorizes the associated *sentiments* into positive, negative, and neutral groups. Positive *opinions* include terms such as "nice", "good", "love", "fantastic", and "really nice", reflecting user satisfaction with specific features. For negative *opinions*, terms like "annoying", "fake", "problem", "not great", and "not so great" highlight common criticisms of certain aspects. Neutral *opinions* feature descriptions like “standard”, “automatic”, “four wheel”, and “SUV”, representing more balanced or informational evaluations.

ASTE techniques often involve complex data preprocessing and feature engineering. According to [8], the most common approaches in the literature are fully supervised, unsupervised, hybrid, semi-supervised, weakly supervised, reinforcement learning, and self-supervised. The main limitation of these techniques is that, for the problem we are addressing, expert information often relies on data that ASTE considers neutral, such as horsepower or battery range in hours. However, these data obviously carry semantic significance within the context of electric car purchases.

Large Language Models (LLMs) demonstrate exceptional potential in Natural Language Processing (NLP), excelling at tasks such as text comprehension, generation, and analysis. Their performance can be further enhanced through prompt engineering, particularly by leveraging techniques like chain-of-thought reasoning, which enables step-by-step problem solving, and few-shot learning, where a minimal number of examples are used to guide the model toward more accurate and context-aware responses [9].

Given the limitations of traditional ASTE techniques in handling technical data within expert reviews, this study explores the integration of LLMs with semantic ASTE to enhance sentiment analysis in electric vehicle evaluations. Specifically, this research aims to answer the following key questions:

1. How can LLMs be adapted to effectively extract and structure both qualitative and quantitative aspects in expert evaluations of electric cars?
2. How can prompt engineering techniques be designed to guide LLMs in performing ASTE effectively, ensuring structured extraction of relevant aspects, opinions, and sentiments on both technical details (normally considered neutral by traditional ASTE) and qualitative sentiment?
3. How can this ASTE-enhanced analysis be systematically integrated into the structured purchasing process?
4. How can sentiment be modeled for better linguistic interpretability without losing precision in calculations?

In pursuit of these objectives, given the current power of LLMs, our hypothesis is that they can be leveraged to develop a semantic ASTE framework that effectively captures the key aspects of electric vehicles in expert evaluations. This approach aims to identify and structure both qualitative sentiment expressions (e.g., interpreting statements like “the power is incredible” as (“power”, “incredible”, “positive”)) and technical or quantitative terms (e.g., processing “the power is 500 HP” as (“power”, “500 HP”, “positive”)). As mentioned, such data are often overlooked or classified as neutral in traditional ASTE approaches. Since defining the ASTE task is a highly formal process, we will rely on Backus–Naur Form (BNF), a standard method for specifying the syntax of formal languages, such as programming languages, which has already been applied in prompt engineering [10], a methodology that could be considered a type of chain-of-thought reasoning. Furthermore, we will employ an extended version of BNF that integrates the semantics of the aspects. To model the extracted sentiments, we will use fuzzy linguistic models, as many authors contend that sentiment is more accurately represented by words than numerical values [11]. This enhanced ASTE process will be integrated into the purchasing process (Figure 1), where expert reviews, particularly those rich in technical details such as electric vehicle evaluations, will serve as the basis for making informed buying decisions.

The rest of the article is structured as follows. In Section 2, we review the state of the art, highlighting the novelty of the ASTE approach we propose. Section 3 provides an overview of the BNF and fuzzy linguistic model fundamentals used in our methodology. In Section 4, we present the proposed model. As a use case, this new vision of ASTE is integrated into the purchasing process illustrated in Figure 1 and applied to the evaluation of several electric car models, focusing on best sellers in the U.S., using expert reviews sourced from YouTube. Finally, in Section 5, we provide conclusions and discuss future work.

2. The State of the Art of ASTE

This state-of-the-art analysis of ASTE has been developed following a methodology inspired by [12]. First, all documents containing keywords related to ASTE were compiled from the Web of Science (Core Collection) on 17 January 2025. Web of Science was chosen for its rigorous indexing standards, ensuring the selection of high-quality, peer-reviewed, and relevant research in the field. The retrieved documents were then verified to ensure they were genuinely related to the technique, resulting in a final selection of 73 articles. To continue with the documentary analysis, SciMAT [13], a widely used tool for bibliometric analysis, was used. The keywords were preprocessed to unify acronyms and ensure terminological consistency in the following steps. The remaining steps of the process were completed, resulting in the strategic diagram displayed in Figure 2, which provides insight into the significance of each theme. This importance is assessed through two metrics: centrality (how interconnected a network is with others) and density (the internal cohesion of the network), as defined by the keywords that characterize the theme within any scientific mapping framework.

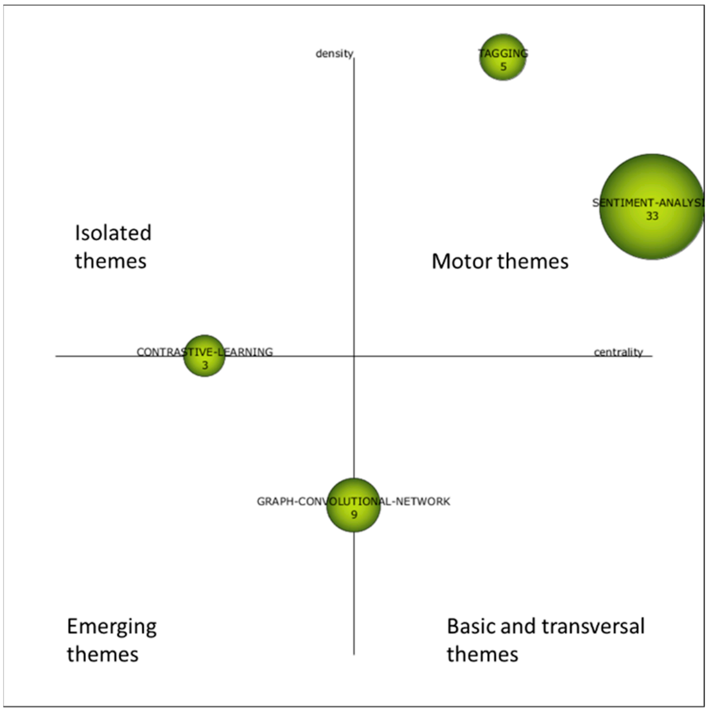

Figure 2. Strategy diagram of ASTE.

Based on this analysis, four fundamental themes were identified that structure the current progress and challenges in ASTE (Figure 3):

- Contrastive learning: Contrastive learning is a self-supervised learning approach that enhances the quality of representations by distinguishing between similar and dissimilar data points. In ASTE, it complements generative models by refining the representations of aspects and opinions, thereby improving overall performance [14,15].
- Graph convolutional network: In ASTE, graph convolutional networks improve the representation of word dependencies and semantic relationships. By integrating

syntactic structures, commonsense knowledge, and other techniques, they enhance triplet detection and contextual understanding, leading to better performance on challenging ASTE tasks [16,17].

- Tagging: In ASTE, tagging assigns labels to words in a sentence to identify aspects, opinions, and sentiment polarities. Traditional tagging models often miss syntactic relationships, leading to errors. Advanced methods, like syntax-aware transformers, enhance tagging by leveraging syntactic dependencies and contextual interactions for more accurate triplet extraction [18,19]. Related to this theme, the use of LLMs for ASTE has been explored in the generation of artificial data (data augmentation) to assist supervised models [20].
- Sentiment analysis: This is a motor area that addresses the simultaneous extraction of aspect terms, opinion terms, and their associated sentiment polarities from text, which is essential for fine-grained sentiment analysis without highlighting the use of any particular technique. According to [8], the most common approaches in the literature for Aspect-Based Sentiment Analysis (ABSA) that include ASTE are fully supervised, which clearly predominates, followed by unsupervised and hybrid as a distant second and third. The remaining approaches, such as semi-supervised, weakly supervised, reinforcement learning, and self-supervised, are residual. For example, the framework applied in [21] uses span-sharing and joint extraction to manage multi-word entities and relations effectively. Another framework [22] refines the extraction process by focusing on the sentiment specific to aspect–opinion pairs, utilizing a position-aware BERT-based approach to enhance accuracy. Additionally, ASTE models have been successfully applied to user reviews, enabling automated feedback analysis and improving user experience [23].

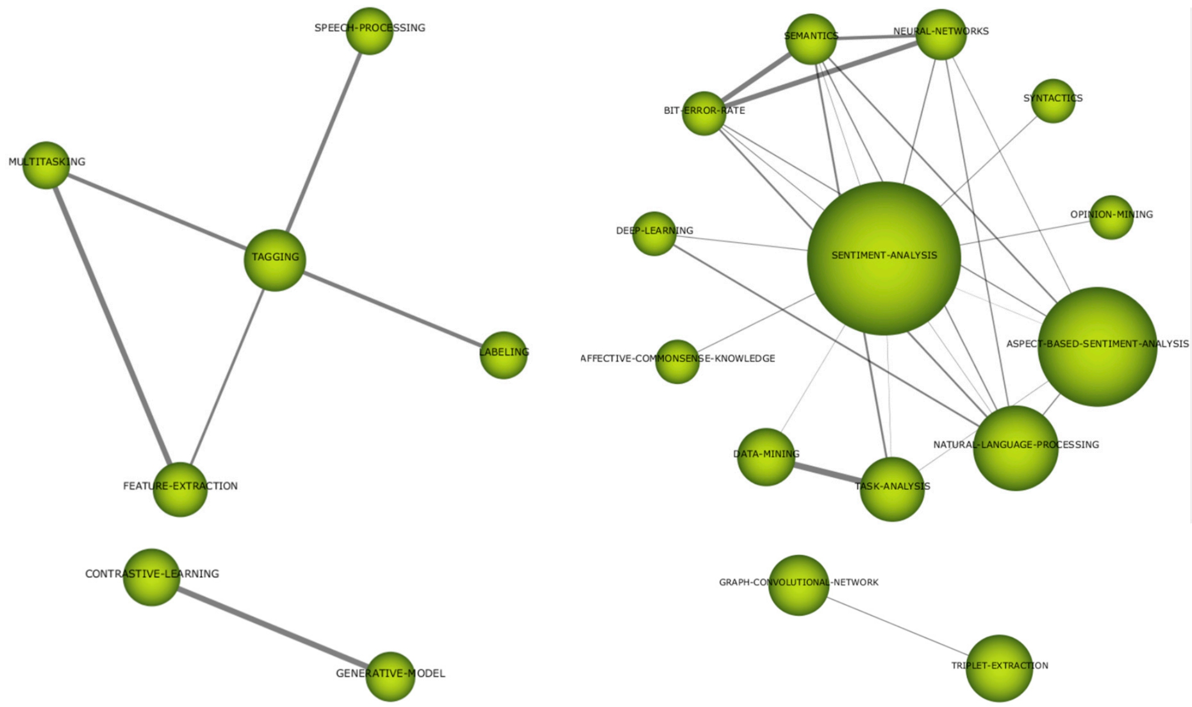

Figure 3. Detail of each theme related to ASTE.

A constant in the ASTE literature is improving syntactic and semantic analysis for a more precise and effective task, as shown by various approaches. In this work, we focus on this task using LLMs by specifying prompts that incorporate a syntactic and semantic specification based on BNF for application to reviews about electric vehicles. These reviews, in addition to subjective opinions, frequently include a multitude of technical data that fall outside the typical ASTE areas studied in the literature. Although, to the best of our knowledge, we have not found any similar models proposed in the literature (extending the search to Scholar and Scopus databases, in addition to Web of Science).

3. Preliminaries

3.1. Context-Free Grammars and Backus–Naur Form

A context-free grammar (CFG) [24] is a mathematical formalism used to describe formal languages through production rules that determine how valid strings are formed in such languages. These grammars are fundamental in formal language theory, programming language design, and syntax analysis. Formally, a CFG is defined as a quadruple $G = <N, \Sigma, P, S>$, where:

- N: A finite set of non-terminal symbols that represent intermediate structures or grammatical categories in the language.
- Σ: A finite set of terminal symbols that are the basic elements of the language (such as characters, words, or tokens).
- P: A finite set of productions, with the form $A \to \alpha$, where:
  - ○ A is a non-terminal symbol ($A \in N$).
  - ○ α is a string composed of terminals and/or non-terminals ($\alpha \in (N \cup \Sigma)^*$). The symbol * indicates zero or more repetitions of the symbols from the set $(N \cup \Sigma)$, allowing for any combination of terminal and non-terminal symbols, including the possibility of an empty string.
- S: The start symbol, an element of N, from which the strings of the language are generated.

They are called "context-free" because the rules of derivation do not depend on the context in which a non-terminal symbol appears. They allow for modeling hierarchical structures, such as mathematical expressions or code blocks, which cannot be represented by regular grammar. BNF is a standard notation used to express CFG [24]. It is commonly used in computer science to specify the syntax of programming languages, offering a compact and readable representation. BNF is specifically designed to describe the rules of a grammar in a clear and precise formal manner. The main elements of BNF notation are as follows:

- Non-terminal symbols represent abstract categories or structures within the language. They are written inside angle brackets (<...>), such as <expression>, <term>, or <factor>.
- Terminal symbols are the basic elements of the language, such as operators (+, *), numbers, or parentheses. They are written exactly as they are, without angle brackets, such as +, *, "(", and "1". Moreover, ε represents the empty string, indicating no symbols in a production.
- Productions define how a non-terminal symbol can expand into a sequence of terminal and/or non-terminal symbols. This indicates that a non-terminal can be replaced by an expression or an alternative.
- For alternatives, the symbol | is used to separate different options in a grammar rule. This means that a non-terminal symbol can expand into one of several possible productions.

As an example, let us consider a simple BNF grammar for arithmetic expressions.

<expression> ::= <expression> "+" <term> | <term>
<term> ::= <term> "*" <factor> | <factor>

<factor> ::= "(" <expression> ")" | <number>
<number> ::= "0" | . . . | "9"

This BNF grammar defines how arithmetic expressions, terms, and factors are structured using terminals like +, *, numbers, and parentheses, and non-terminals like <expression>, <term>, and <factor>.

BNF, originally designed to describe the syntax of formal languages, can be extended through attribute grammars to incorporate semantics [25]. These grammars allow attributes to be associated with non-terminal symbols and define rules to compute their values, providing meaning beyond the syntactic structure. For example, the rule {<A01>.meaning = "Price represents the cost of the electric vehicle."} assigns a synthetic attribute .meaning to the symbol <A01>, which semantically describes the aspect, enabling the contextual analysis and processing of the language.

3.2. Definition of Fuzzy Linguistic Variable Using the 2-Tuple Linguistic Model

A fuzzy linguistic variable is a concept used to represent qualitative information quantitatively through linguistic terms that are associated with a degree of membership within a continuous interval. According to [26], a linguistic variable X is characterized by the following five elements $<L, S, UD, G, M>$, where:

1. L: The name of the linguistic variable.
2. S: A finite set of linguistic terms, $S = \{s_0, \ldots, s_g\}$, representing different values of the variable, such as "very low", "low", "medium", "high", etc.
3. UD: A universe of discourse, defined as a closed interval of real numbers.
4. G: A syntactic rule that generates the linguistic terms.
5. M: A semantic rule that assigns meaning to each linguistic term through a compatibility function $f(c){:}UD \rightarrow [0{,}1]$, which evaluates the membership degree of a value in UD to the linguistic term.

The fuzzy linguistic 2-tuple [27] introduces a mechanism for representing intermediate values between linguistic terms in a precise and interpretable manner. Let $\beta \in [0,g]$ represent the result of a symbolic aggregation operation on the indexes of terms from the linguistic term set S, where $\beta \notin \{0, \ldots, g\}$. The model represents β using ordered pairs (s_i, α), and is defined as follows:

- $s_i \in S$: The linguistic term in S closest to β, determined by i = round(β).
- $\alpha \in [-0.5, 0.5)$: A numerical value that represents the symbolic deviation or distance of β from the center of s_i.

To perform calculations, the model defines the following bidirectional transformation functions:

- From numerical value to 2-tuple: $\Delta(\beta) = (s_i, \alpha)$,where i = round(β), $\alpha = \beta - i$. When $\alpha = 0$, we consider $\Delta(\beta) = s_i$.
- From 2-tuple to numerical value: $\Delta^{-1}(s_i, \alpha) = i + \alpha = \beta$.

We can redefine the fuzzy linguistic variable X integrating the principles of the 2-tuple linguistic model. Thus, a 2-tuple linguistic variable X is characterized by the following elements:

1. L: The name of the linguistic variable (e.g., "Satisfaction").
2. S: In this work, we consider as a set an odd set of symmetric linguistic terms, $\{s_0, s_1, \ldots, s_g\}$, where:
    - The terms are evenly distributed around a central term (e.g., $s_{g/2}$ = "medium").
    - The rest of terms represent a qualitative value and are distributed symmetrically around the central term (e.g., s_g = "very low", s_g = "very high").

3. *UD*: The universe of discourse, defined typically as [0,1], mapping quantitative inputs to the linguistic domain.
4. *G*: is the function that generates a 2-tuple from a numerical value $\beta \in UD$, using the transformation function **Δ**: $G = \Delta(\beta) = (s,\alpha)$, with $s \in S$. We can express *G* using BNF:
   <2-tuple> ::= “(“ <term> “,“ <alpha> ”)” | <term>
   <term> ::= “s_1” | . . . | “s_g”
   <alpha> ::= <real> | “0”
   <real> ::= <positive-real> | <negative-real>
   <positive-real> ::= <digits> “.” <fraction> | “0” “.” <fraction>
   <negative-real> ::= “-” <digits> “.” <fraction>
   <digits> ::= “0” | . . . | “5”
   <fraction> ::= <digit> <digit-seq>?
   <digit> ::= “0” | . . . | “9”
   <digit-seq> ::= <digit> <digit-seq>*
   If <alpha> is zero, we represent the 2-tuple value with only the corresponding <term>.
5. *M:* A semantic rule that assigns meaning to each 2-tuple (s_i, α) using a triangular membership function $f(s_i, \alpha)$. This function is defined to complete the following:
   - Center the triangle at $\Delta^{-1}(s_i, 0) = i$ (the index of s_i).
   - Adjust the triangle's width based on the number of terms g to ensure the universe of discourse is covered.
     Let X be a 2-tuple linguistic variable characterized by the following instances: $X_1 = (s_1, \alpha_1), \ldots, X_k = (s_k, \alpha_k)$. We can define different means of these instances as follows:
   - Arithmetic mean:

$$\overline{X} = \Delta\left(\frac{1}{k}\sum_{i=1}^{k} \beta i\right), \beta i = \Delta^{-1}(s_i, \alpha_i) \tag{1}$$

   - Weighted mean based on a vector of weights, $W = [w_1, \ldots, w_k]$, associated with each instance X_i, where $w_i \geq 0$ and $\sum_{i=1}^{k} w_i = 1$, is defined as follows.

$$\overline{X}_w = \Delta\left(\sum_{i=1}^{k} w_i \beta i\right), \beta i = \Delta^{-1}(s_i, \alpha_i) \tag{2}$$

3.3. Aspect Sentiment Triplet Extraction

Aspect Sentiment Extraction (ASE) is a fundamental task in Natural Language Processing that focuses on identifying specific aspects mentioned in a text and determining the sentiment expressed toward each aspect. For instance, in a product review such as “The range of this electric car is fantastic, but the charging speed is disappointing”, ASE identifies the aspects (“range” and “charging speed”) and their corresponding sentiments (positive for “range” and negative for “charging speed”). ASE helps extract structured information from unstructured text, making it invaluable for applications like customer feedback analysis and opinion mining. Although ASE provides these useful insights by associating aspects with sentiments, it does not capture the specific opinion terms or phrases that justify the sentiment. This limitation is addressed by ASTE, which adds a third dimension (opinion terms) into the analysis. This provides a richer understanding of the text by identifying not only what is being discussed and how it is perceived, but also why the sentiment is expressed. We can define ASTE as the following function:

$$f_{ASTE}: D \rightarrow \{(a_i, o_i, t_i) \mid a_i \in A, o_i \in O, t_i \in T\}$$

where:

1. D is the set of input documents or sentences;
2. A is the set of possible aspects;
3. O is the set of opinion terms explicitly expressing sentiments about the aspects;
4. T is the set of sentiment labels, typically {negative, neutral, positive}.

Thus, in ASTE, each triplet (a_i, o_i, t_i) associates an aspect a_i with its opinion o_i and its sentiment t_i. For instance, if we have the input text $d \in D$ such that d = "The range of this electric car is fantastic, but the charging speed is disappointing, and the interior feels luxurious.", the output would be as follows.

$f_{ASTE}(d)$ = {("range", "fantastic", "positive"), ("charging speed", "disappointing", "negative"), ("interior", "luxurious", "positive")}.

4. Proposed Model

In this section, we propose a system that enhances electric vehicle selection by addressing key limitations of traditional ASTE methods. Our approach is differential due to the following aspects:

- Combination of LLMs and BNF for structured ASTE: Traditional ASTE methods often classify technical data as neutral, limiting their effectiveness. Our approach leverages LLMS to extract both qualitative sentiment expressions and technical quantitative aspects while incorporating BNF to enhance syntactic and semantic precision, ensuring a more structured and comprehensive analysis.
- Application of fuzzy linguistic models: Instead of relying solely on numerical sentiment scores, we employ fuzzy linguistic models to better represent the gradual nature of sentiment, improving interpretability.
- Integration into the purchasing process: Unlike prior ASTE studies that focus on generic sentiment extraction, our method aligns with the structured purchasing framework (Figure 1), ensuring its applicability in real-world decision making.

This model ensures a more structured, interpretable, and practical application of ASTE for informed electric vehicle selection, as detailed below.

4.1. Problem Recognition

In this stage, the buyer identifies a need or problem that requires a solution. During this phase, they may also establish criteria or priorities that will guide their decision, such as budget, quality, or specific features. Additionally, the buyer selects initial alternatives that they believe may fulfill their needs. These criteria and alternatives can evolve as they progress through the process, and, in some cases, the buyer may revisit and adjust them as new information is gathered or further options are considered. In the specific problem we are dealing with, i.e., electric car selection, in the problem recognition stage, the buyer identifies their need for an electric car and begins to explore options that may fulfill this need. This phase involves understanding the key aspects of electric cars or criteria, such as performance, range, design, and price, and forming an initial set of alternatives to evaluate.

The use of ASTE is particularly justified in this context because these key aspects of electric cars are frequently evaluated by experts through comprehensive reviews and analyses. These evaluations often include detailed examinations of specific attributes, accompanied by explicit opinions and sentiment expressions regarding each aspect. This process inherently involves identifying relevant aspects (e.g., "battery efficiency"), extracting corresponding opinion terms (e.g., "excellent"), and classifying sentiments (e.g., positive). By applying ASTE, we can formalize this evaluation process, enabling the structured analysis of expert reviews and related documents. This structured approach supports informed and rational decision making, particularly when selecting among competing electric car options.

Based on this justification, we propose modeling the evaluation of electric cars within the framework of ASTE. Let $C = \{C_1, \ldots, C_{\#C}\}$ represent the set of electric cars that the buyer is considering, where $\#C$ is the number of cars to be considered. For each car C_k, we define the following ASTE function:

$$f^k{}_{ASTE}: D^k \rightarrow \{(a_i, o_i, t_i) \mid a_i \in A, o_i \in O^k, t_i \in T\}$$

where:

1. D^k: This represents the set of input documents or sentences specifically related to car C^k, such as expert reviews or opinions.
2. A: We define the set of possible aspects of interest to the buyer as $A = \{A_1, \ldots, A_{\#A}\}$, where $\#A$ is the number of aspect related to the electric car to be considered by the buyer. Associated to this set, we will define the set with detailed descriptions of each of these aspects $AD = \{AD_1, \ldots, AD_{\#A}\}$. In this paper, we will consider the aspects shown in Table 1 based on [2].
3. O^k: This is the set of opinion terms found in D^k that explicitly express sentiments about the aspects of car C^k.
4. T: We consider the set of sentiment labels defined as T = {negative, neutral, positive}.

Table 1. Aspects to be considered in the electric car buying process.

Aspect $\boldsymbol{A_i}$	**Aspect Description** $\boldsymbol{AD_i}$
Price	Electric car prices vary depending on the model and features. Prices under EUR 40,000 are considered positive; those between EUR 40,000 and EUR 55,000 (inclusive) are neutral; and prices higher than EUR 55,000 are viewed as negative.
Acceleration	Acceleration time measures how quickly the car goes from 0 to 100 km/h (or 0 to 60 mph). Under 5 s is positive, 5–8 s is neutral, and over 9 s is negative, especially for premium models.
Battery range	The maximum distance a car can travel on a full charge. Over 480 km (300 miles) is positive. A range of 400–480 km (250–300 miles) is neutral. Under 400 km (250 miles) is negative for long trips.
Full charge time	The time required to charge a battery from 0% to 100% using a standard charger, excluding fast charging. Under 8 h is positive, 8–12 h is neutral, and over 12 h is negative.
Fast charge time to 80%	Under 45 min is positive, 45–60 min is neutral, and over 60 min is negative.
Top speed	Over 200 km/h (125 mph) is positive, 150–200 km/h (93–125 mph) is neutral, and under 150 km/h (93 mph) is negative for highways or performance cars.
Energy usage	Energy consumption per 100 km (or 62 miles). Under 15 kWh is positive, 15–20 kWh is neutral, and over 20 kWh is negative.
Accumulator capacity	Refers to energy storage, measured in kWh. Over 75 kWh is positive, 60–75 kWh is neutral, and below 60 kWh is negative for long trips or high energy use.
Engine power	Engine power, measured in horsepower (HP). Less than 250 HP is negative, 250–300 HP is neutral, and over 300 HP is positive.
Vehicle Weight	Under 1800 kg is positive, 1800–2200 kg is neutral, and over 2200 kg is negative for range and handling.
Maximum Torque	Above 500 Nm is positive, 300–500 Nm is neutral, and below 300 Nm is negative for acceleration and towing.

Based on [28], we will select three of the top best-selling electric vehicles in the United States in 2024 as a case study in this article, ordered by sales as follows: C_1 = "Tesla Model 3", C_2 = "Tesla Model Y", and C_3 = Hyundai Ioniq 5".

4.2. Information Search

In this step, the buyer gathers information about possible solutions, exploring products, services, or brands that could fulfill their needs. This includes consulting online resources, reviews, recommendations, or direct experiences. In some cases, after obtaining the necessary information, the buyer may return to the previous phase, redefining criteria and/or alternatives, adjusting expectations, etc.

In our case, in this phase, and following the schema defined as ${f^k}_{ASTE}$, we will extract the expert opinions, represented as D^k. This extracted information will be essential for guiding the selection of an electric vehicle, providing valuable insights into the strengths and weaknesses of each model. As mentioned above, we will use YouTube as the basic source of information. Based on the selected models analyzed as our case study, C_k, the YouTube channel with the most views that includes analyses of these models is ***Carwow*** [29]. This channel offers detailed reviews, comparisons, and performance tests of the vehicles in question. In Table 2, the specific videos chosen for our analysis are detailed. To obtain the expert review text from each video, represented as $\boldsymbol{D^k}$, we use the *youtube_transcript_api* Python library [30] to transcribe the content. This transcription process converts spoken content into text, creating a dataset suitable for analysis. Additionally, Table 2 includes the first sentences from each transcription, providing an initial overview of the extracted content.

Table 2. YouTube videos selected in information search stage.

$\boldsymbol{C_k}$	Ref.	Views	Duration	$\boldsymbol{D^k}$
C_1	[31]	2.2 M	14 m 14 s	This is the tesla model 3 and it's a little bit like a one plus mobile phone in the way it has almost the same performance as the more expensive model s [. . .]
C_2	[32]	9.0 M	17 m 46 s	This is the new tesla model y and it is effectively a tesla model 3 that's been like stretched in photoshop now really look i've stretched the model 3 in photoshop [. . .]
C_3	[33]	1.9 M	15 m 12 s	This is the new high andai ironic 5 and it's a little bit like US Open tennis champion Emma ranu because it's one a grand slam smash of a car it's [. . .]

4.3. Evaluation of Alternatives

In this phase, the buyer compares different options based on the criteria established earlier. If the available options do not meet their expectations, they may return to earlier phases to refine their criteria or seek additional information. Continuing with our case, in this phase, our objective is to obtain experts' valuations on the different aspects of the cars and then determine the final valuation of each car.

4.3.1. Extraction of Experts' Valuations of Each Car

In this stage, we extract experts' evaluations of the key aspects, A, defined in Table 1, for each electric car considered in our study. These evaluations are essential for comparing the cars' performance and determining their suitability based on the buyer's criteria. To ensure a structured and precise analysis, we use the ASTE function described earlier, ${f^k}_{ASTE}$, applied to the textual data, D^k, obtained during the information search phase. This involves processing these expert reviews to identify the opinion terms (O_k) and sentiment labels (T) associated with the aspects (A) defined with its semantics (AD) for each car (C_k).

As mentioned, to effectively analyze expert reviews and extract evaluations, we employ LLMs as an alternative to traditional approaches. Our LLM-based approach is divided into two phases.

1. **Phase: Setup**

This phase establishes the general LLM’s behavior, parameters, and tasks to ensure accurate and reliable performance while addressing common issues associated with LLMs, such as hallucinations, irrelevant outputs, or inconsistencies. In this work, we use the *gemini-1.5-pro* model using the API of the Google Generative AI library in Python. The specific details we have used in our use case are shown in Table 3. These are explained in more detail below.

Table 3. Setup of LLM for ASTE.

<table>
<tr><th>Sub-Phase</th><th></th></tr>
<tr><td>Model configuration</td><td>System instruction: Act as an analyst specializing in extracting Aspect, Sentiment, and Opinion Triplets (ASTE) from opinions on electric cars. Strictly adhere to the rules defined in the prompt and apply BNF grammar and its semantic definitions precisely. Focus on the semantic ranges of the aspects to classify sentiment accurately as positive, negative, or neutral. Produce the results in the specified CSV format and rigorously follow the BNF rules and semantics to ensure strict compliance with the framework.
LLM parameters: temperature = 0.4, top_p = 0.9, top_k = 600</td></tr>
<tr><td>Task definition through prompting</td><td>LLM prompt:
INSTRUCTIONS
This task has two phases:
1. **Setup Phase**: Define the CONTEXT, OUTPUT FORMAT, GRAMMAR IN BNF WITH SEMANTIC, FEW-SHOT EXAMPLES, and EXCEPTIONS to establish the task rules.
2. **Execution Phase**: The EXECUTION PHASE process several inputs text (<phrase>) provided in subsequent calls (prompts) and generate the corresponding output format.
CONTEXT
Extract aspect sentiment triplets (<aste>) from input text (<phrase>). Each triplet (<triplet>) represents:
- <aspect>: The attribute being discussed.
- <opinion>: The related opinion or descriptive phrase.
- <sentiment>: The expressed sentiment.
Apply the “meaning” attribute of each aspect to extract the rules accurately and ensure consistency and clarity in the extraction process. Use the defined intervals within the “meaning” attribute to guide the classification and extraction of aspects.
OUTPUT FORMAT
CSV format:
- Fields enclosed in double quotes (‘“’) and separated by comma (‘,’).
- If no triplets are found, return a CSV with headers only: “aspect”, “opinion”, “sentiment”.
GRAMMAR IN BNF WITH SEMANTIC
<aste> ::= <triplet> | <triplet> “,” <aste>
<triplet> ::= “(“ <aspect> “, “ <opinion> “, “ <sentiment> ”)”
<aspect> ::= <A01> | <A02> | <A03> | <A04> | <A05> | <A06> | <A07> | <A08> | <A09> | <A10> | <A11>
<opinion> ::= <phrase>
<sentiment> ::= “positive” | “negative” | “neutral”
<phrase> ::= <word> | <word> “ ” <phrase>
<word> ::= <character> | <character> <word>
<character> ::= “a” | “b” | “c” | . . . | “z” | “A” | “B” | “C” | . . . | “Z” | “0” | “1” | . . . | “9” | “.” | “!”
<A01> ::= A_1 { <A01>.meaning = AD_1}
. . .
<A11> ::= A_{11} { <A11>.meaning = AD_{11}}
FEW-SHOT EXAMPLES
Input (<phrase>):</td></tr>
</table>

Table 3. *Cont.*

Sub-Phase	
Task definition through prompting	<phrase> ::= "The price of the new EV is a hefty $60,000. Its acceleration is decent, reaching 0 to 60 mph in 7 s. The battery provides a good range of 450 km. However, a full charge takes a whopping 14 h. The fast charge to 80% is completed in just 30 min. The top speed is an impressive 220 km/h. Energy consumption is around 18 kWh/100 km. The accumulator capacity is 80 kWh. It boasts 400 HP, but it weighs 2500 kg. Maximum torque is 450 Nm." **Output (<aste> in CSV Format):** ```csv "aspect", "opinion", "sentiment" "A01(Price)", "a hefty $60,000", "negative" \n "A02(Acceleration)", "decent, reaching 0 to 60 mph in 7 s", "neutral" \n "A03(Battery range)", "a good range of 450 km", "neutral" \n "A04(Full charge time)", "a whopping 14 h", "negative" \n "A05(Fast charge time to 80%)", "just 30 min", "positive" \n "A06(Top speed)", "an impressive 220 km/h", "positive" \n "A07(Energy usage)", "around 18 kWh/100 km", "neutral" \n "A08(Accumulator capacity)", "80 kWh", "positive" \n "A09(Engine Power)", "400 HP", "positive" \n "A10(Vehicle weight)", "2500 kg", "negative" \n "A11(Maximum torque)", "450 Nm", "neutral"\n ``` # EXCEPTIONS If <aste> is not found, generate a CSV with the headers only.``` **LLM output:** Okay, I'm ready for the Execution Phase. Please provide the input '<phrase>'.

- **Model configuration**. On the one hand, we specify the directive *system instruction*, in which we clearly define the behavior of the model as an analyst specialized in ASTE from opinions on electric cars. On the other hand, the typical model parameters are carefully chosen to balance the precision and flexibility of the outputs. Namely, a low *temperature* (close to 0) reduces randomness, ensuring deterministic and reliable responses that avoid overly creative or divergent answers, which are unsuitable for structured tasks like our case. A high *top-p* value (close to 1) narrows the focus to the most relevant outputs, minimizing the risk of hallucinations while maintaining coherence. Additionally, the *top-k* value provides the model with enough options to handle nuanced language and complex inputs without compromising accuracy.
- **Task definition through prompting**. The first prompt sent to the model defines its tasks with precision, ensuring clear alignment with the ASTE extraction objective (i.e., the $f^k{}_{ASTE}$ function). It uses a structured approach, where the semantic BNF grammar can be considered an ASTE technique (<aste>) within the scope of chain-of-thought reasoning, as it clearly establishes the step-by-step process for extracting triplets (<triplet>, i.e., (a_{i}, o_{i}, t_i)), enforcing syntactic rules and integrating the semantic "meaning" attributes (AD_i) of aspects (A_i) to guide accurate extraction. In this BNF definition, the input of the ASTE process (i.e., <phrase>, i.e., D^k) is not specified as it will be provided in the next stage. However, the prompt includes few-shot examples, where the input is provided (<phrase>) to demonstrate the expected input–output relationship. This combination, through the application of the semantic BNF grammar, ensures the model processes inputs systematically, producing structured CSV outputs while handling exceptions predictably.

2. **Phase: Execution**

For each car (C_k), a different prompt is executed, providing the BNF grammar with the input (<phrase>), i.e., the expert's opinion contained in the textual data (D^k). The LLM returns the output (<triplet>, i.e., (a_{i}, o_{i}, t_i)) in the specified CSV format. Each prompt call includes the current input and the previous conversation to ensure the LLM maintains context. The details of this process are shown in Table 4.

Table 4. Execution of LLM for ASTE.

C_k	LLM Prompt (D^k)	LLM Output ($f^k{}_{ASTE}$ (D^k))
C_1	# EXECUTION PHRASE 1 <phrase> ::= D^1	```csv "aspect", "opinion", "sentiment" "A01(Price)", "35,000 (dollars in America)", "positive" "A01(Price)", "50,000 euros (Europe)", "neutral" "A01(Price)", "50,000 pounds (UK-estimated)", "neutral" "A09(Engine Power)", "350 horsepower (long-range version)", "positive" "A09(Engine Power)", "450 horsepower (performance version)", "positive" "A03(Battery range)", "up to 300 miles (with 75 kilowatt hour battery)", "neutral" "A04(Full charge time)", "a day and a half (from normal three-pin socket)", "negative" "A04(Full charge time)", "half a day (with special fast AC charger)", "negative" "A05(Fast charge time to 80%)", "about an hour (from a 50 kilowatt charger)", "negative" "A05(Fast charge time to 80%)", "about 35 min (from a Tesla Supercharger)", "positive" "A05(Fast charge time to 80%)", "75 miles of range in just five minutes (with new V3 Supercharger)", "positive" "A02(Acceleration)", "0 to 60 mph in 3.6 s", "positive" ```
C_2	# EXECUTION PHRASE 2 <phrase> ::= D^2	```csv "aspect", "opinion", "sentiment" "A01(Price)", "55 000 pounds", "neutral" "A10(Vehicle weight)", "taller and heavier (than Model 3)", "negative" "A02(Acceleration)", "0–60 mph in 4.8 s (long range version)", "positive" "A02(Acceleration)", "0–60 mph in 3.5 s (performance version-claimed)", "positive" "A09(Engine Power)", "394 horsepower (long range version, dual motor)", "positive" "A09(Engine Power)", "450 horsepower (performance version, dual motor)", "positive" "A08(Accumulator capacity)", "70 kilowatt hour battery pack", "neutral" "A03(Battery range)", "315 miles maximum", "positive" ```
C_3	# EXECUTION PHRASE 3 <phrase> ::= D^3	```csv "aspect", "opinion", "sentiment" "A01(Price)", "£37,000 (starting price)", "positive" "A09(Engine Power)", "170 horsepower (single motor, 58 kWh battery)", "negative" "A09(Engine Power)", "217 horsepower (single motor, 73 kWh battery)", "negative" "A09(Engine Power)", "355 horsepower (dual motor, 73 kWh battery)", "positive" "A08(Accumulator capacity)", "58 kWh battery", "negative" "A08(Accumulator capacity)", "73 kWh battery", "neutral" "A03(Battery range)", "240–300 miles (depending on version)", "neutral" "A05(Fast charge time to 80%)", "20 min (from 10% to 80% with 220 kW fast charging)", "positive" "A02(Acceleration)", "0–60 mph in 5.2 s (dual motor-claimed)", "neutral" "A02(Acceleration)", "0–60 mph in 4.79 s (dual motor-tested)", "positive" ```

4.3.2. Determination of Final Valuation of Each Car

In this stage, for each car (C_k) and each aspect ($A_r \in A$), we define a 2-tuple linguistic variable D_r^k, characterized by $S = \{s_0 = \text{negative}, s_1 = \text{neutral}, s_2 = \text{positive}\}$, where each label is defined using its corresponding triangular membership function. Consequently, the values of this variable are the 2-tuples (t_i, 0), corresponding to all triplets i of the given car and aspect $a_i = A_r$, i.e., (a_i, o_i, t_i).

To calculate the evaluation for each aspect and each car, we compute the arithmetic mean of the variable D_r^k, denoted as $\overline{D_r^k}$, using Equation (1). For those aspects for which no valuation has been found, we consider their value as neutral. Finally, for each car, the average values of each aspect ($\overline{D_r^k}$) are aggregated, taking into account the weights assigned to each aspect by experts according to the article [2], W = (0.12, 0.07, 0.05, 0.08, 0.15, 0.04, 0.02, 0.11, 0.18, 0.05, 0.14), using the weighted mean as defined in Equation (2), resulting in the corresponding final values $\overline{D_W^k}$. The results of this phase are shown in Table 5.

Table 5. Final valuation of each car.

C_k	$\overline{D_1^k}$	$\overline{D_2^k}$	$\overline{D_3^k}$	$\overline{D_4^k}$	$\overline{D_5^k}$	$\overline{D_6^k}$	$\overline{D_7^k}$	$\overline{D_8^k}$	$\overline{D_9^k}$	$\overline{D_{10}^k}$	$\overline{D_{11}^k}$	$\overline{D_W^k}$
C_1	$(s_1, 0.167)$	s_2	s_1	s_0	$(s_1, 0.167)$	s_1	s_1	s_1	s_2	s_1	s_1	$(s_1, 0.133)$
C_2	s_1	s_2	s_2	s_1	s_1	s_1	s_1	s_1	s_2	s_0	s_1	$(s_1, 0.129)$
C_3	s_2	$(s_2, -0.25)$	s_1	s_1	s_2	s_1	s_1	$(s_0, 0.25)$	$(s_1, -0.167)$	s_1	s_1	$(s_1, 0.092)$

4.4. Purchase Decision

After sufficient evaluation, the buyer selects the option that best aligns with their criteria and proceeds with the purchase. Following with our case, this decision could be based on Table 5. It is important to note that each user may have different priorities compared to those used in our study to calculate the final score for each car. Therefore, the weights vector (W) would need to be obtained to reflect the specific preferences of the individual buyer and then the global score of each car ($\overline{D_r^k}$) would need to be obtained again. This customization ensures that the evaluation aligns with what the buyer values most, such as price, performance, or charging time. In addition to the overall score ($\overline{D_r^k}$), the final decision could also take into account the individual values of each aspect ($\overline{D_r^k}$) and their associated triplets (a_i, o_i, t_i) (see Table 4). This allows for a more qualitative analysis, enabling the buyer to focus on specific strengths or weaknesses of each car and aspect.

Finally, in the specific use case analyzed, and with the weights specified by the experts, the three cars performed almost identically in terms of their overall scores, as reflected in Table 5. Given this near-tie, the potential buyer should consider this carefully, focusing on the aspects and triplets most relevant to their preferences to make a final decision that best suits their individual needs.

4.5. Post-Purchase Behavior

Following the purchase, the buyer reflects on their satisfaction. Positive experiences can lead to loyalty and future purchases, while dissatisfaction may prompt the reconsideration of the process or brand. In the context of our model, post-purchase opinions and reviews, especially those shared on social media and online platforms, are crucial as they could be incorporated as input into our model. Over time, this feedback loop would improve the model's accuracy, making it increasingly valuable for guiding future buyers.

5. Discussion

In the applied use case, three vehicles were analyzed, selected based on their status as top-selling electric vehicles and the high number of views on the YouTube videos used for sentiment analysis. This selection ensures that the vehicles analyzed are both market-relevant and widely discussed among consumers, reinforcing the applicability of our findings, with their results presented in Table 5. The model demonstrated an advanced ability to differentiate and contextualize evaluations between different versions of the same vehicle, enabling a more precise and detailed analysis.

For example, the C_1 model received a positive evaluation regarding engine power, with 350 HP in its standard version and 450 HP in its high-performance version. However, its full charging time from a household outlet was rated negatively, as it takes approximately a day and a half, representing a competitive disadvantage in terms of convenience.

The C_2 model, on the other hand, received favorable comments regarding its acceleration and engine power, with 394 HP in its long-range version and 450 HP in its high-performance version. Nevertheless, its high weight compared to similar models was identified as a negative aspect, as it impacts the vehicle's efficiency and maneuverability.

Finally, the C_3 model received mixed reviews. Its starting price of GBP 37,000 was well received, positioning it as an accessible option within its segment. However, the power of its basic versions (170 HP and 217 HP) generated less favorable opinions, reflecting a perception of limited performance. In contrast, the high-performance version, with 355 HP, was positively evaluated, as was its fast-charging capability, which allows it to reach 80% in just 20 min, establishing itself as a key competitive advantage.

It is worth noting that the final evaluation obtained through ASTE analysis with LLMs matches the sales ranking of the three vehicles evaluated in the U.S. This suggests that the extracted results from reviews correlate with consumers' purchasing decisions, validating the effectiveness of the proposed model in interpreting automotive market preferences.

The integration of LLMs into the ASTE process has significantly improved the efficiency of vehicle analysis. Although ASTE has been frequently used in this context, its performance has been limited due to the technical nature of many reviews, where traditional methods classify numerous attributes as neutral without extracting relevant information. In this study, the combination of LLMs with a BNF semantic extension and their integration into best practices of prompt engineering (chain-of-thought and few-shot learning) optimized data interpretation. This semantic extension allowed the LLM to achieve a deeper understanding of each evaluated element, eliminating the need for iterative preprocessing and reducing reliance on complementary analyses such as graph visualization.

Once the ASTE triplets were extracted, sentiment representation was carried out using a 2-tuple linguistic model, allowing the word aggregation process to occur without loss of information. This approach enabled the refinement of various sentiments through semantic displacement, better capturing the nuances of each vehicle's evaluations and providing a more faithful representation of expert assessments.

LLMs continue to evolve, and, in this study, various versions were tested to assess their effectiveness in extracting technical information. In particular, lighter models such as *flash* did not achieve satisfactory results in this type of analysis, whereas more recent versions showed better performance. One of the most significant challenges in using these models was the formulation of effective prompts. Our approach addresses this issue by employing a formal language based on semantic grammars, allowing models to accurately understand and process technical data without relying on ambiguous or inconsistent formulations.

While LLMs are capable of interpreting technical information, this study highlights the importance of providing them with a well-structured expert description to ensure that results are more reliable and reproducible. In this regard, our approach not only optimizes information extraction, but also ensures that the interpretation of results reflects the knowledge and expertise of subject matter specialists. This makes the system more robust and adaptable to technical evaluation contexts, expanding its applicability and reliability across different domains.

6. Conclusions and Future Work

This study demonstrated how the combination of ASTE with LLMs and semantic BNF grammar can significantly enhance decision support in the general selection of electric vehicles, regardless of brand, model, or market conditions. By leveraging expert reviews expressed in video and structuring them into interpretable insights, our approach overcomes key limitations of traditional sentiment analysis and decision support models by effectively integrating both quantitative and qualitative data. As shown in the discussion, the ASTE results closely align with actual sales rankings of analyzed vehicles, reinforcing the model's practical impact in reflecting consumer preferences.

This challenge is prevalent in diverse sectors such as tourism and financial investments, where decision making requires balancing qualitative and quantitative factors. For instance, in tourism, customers assess both measurable aspects like distance to main attractions (e.g., distance to the beach), official star rating, number of beds per room, and availability of essential facilities (e.g., presence of a gym, parking, and restaurant), as well as subjective elements such as customer satisfaction scores and user-generated ratings on platforms like TripAdvisor. Likewise, in stock market investments, technical analysis provides numerical insights, while fundamental analysis evaluates qualitative factors such as market trends and company stability. Both sets of data are crucial in providing a balanced and realistic view in decision making.

Therefore, looking ahead, this work envisions a scenario in which purchasing processes, in general, evolve into an automated decision model. This process can be automated through various AI agents responsible for managing each phase of the process described in Section 1. Some agents will assist in precisely defining the purchase need and some will focus on gathering information (which, as demonstrated in this study, can come from any reliable source and be in any format, be it video, reports, etc.), while others will evaluate the available alternatives. There will also be agents responsible for the final purchasing decision, as well as agents focused on post-purchase actions, such as verifying the quality of the purchase or managing reviews. The core of these agents will be AI-based natural language models capable of interpreting and responding flexibly to user needs (even for non-technical users). However, these interactions may exhibit unpredictable behaviors. Therefore, the use of formal languages, such as BNF with semantic extensions, as presented in this study, will allow for precise specifications, ensuring more reliable and structured processes.

With these improvements, the automation of complex decision making will become increasingly precise, scalable, and personalized, marking a significant advancement in the use of AI for consumer guidance. Consequently, the complete automation of this process, its comparison with other decision-making models for electric vehicle purchasing, its application to other domains, and the integration of various AI agents to manage different phases of the decision process represent key advancements that define the future of this work.

Author Contributions: Conceptualization, Á.C.-A., M.M.C.-M. and M.C.P.-M.; methodology, Á.C.-A., M.M.C.-M. and M.C.P.-M.; software, Á.C.-A.; validation, Á.C.-A., M.M.C.-M., M.C.P.-M. and M.S.-R.; formal analysis, Á.C.-A., M.M.C.-M., M.C.P.-M. and M.S.-R.; investigation, Á.C.-A., M.M.C.-M., M.C.P.-M. and M.S.-R.; resources, Á.C.-A.; data curation, Á.C.-A.; writing—original draft preparation, Á.C.-A.; writing—review and editing, Á.C.-A., M.M.C.-M., M.C.P.-M. and M.S.-R.; visualization, Á.C.-A.; supervision M.M.C.-M. and M.C.P.-M.; project administration, M.M.C.-M. and M.C.P.-M.;

funding acquisition, Á.C.-A., M.M.C.-M., M.C.P.-M. and M.S.-R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Complutense University Research Project titled "Impacts of Artificial Intelligence on the Tourism Industry from a Social Science Perspective" (project number 506-2024), as well as the Telefónica Chair on Smart Cities at Universitat Rovira i Virgili and Universitat de Barcelona (project number 42.DB.00.18.00).

Data Availability Statement: The data that support the findings of this study are openly available on YouTube at https://www.youtube.com/ (accessed on 29 December 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Haustein, S.; Jensen, A.F. Factors of electric vehicle adoption: A comparison of conventional and electric car users based on an extended theory of planned behavior. *Int. J. Sustain. Transp.* **2018**, *12*, 484–496. [CrossRef]
2. Więckowski, J.; Białas, P.; Cheba, K.; Wątróbski, J.; Jankowski, P.; Ziemba, P. Complex sensitivity analysis in multi-criteria decision analysis: An application to the selection of an electric car. *J. Clean. Prod.* **2023**, *390*, 136051. [CrossRef]
3. Wijaya, A.; Tj, H.W.; Wahyoedi, S. The effect of car sales e-commerce platform on new car buying decisions mediated by digital advertising media and YouTube media. *Daengku: J. Humanit. Soc. Sci. Innov.* **2022**, *2*, 596–607. [CrossRef]
4. Kotler, P.; Burton, S.; Deans, K.; Brown, L.; Armstrong, G. *Marketing*; Pearson Higher Education: London, UK, 2017.
5. Dahiya, R.; Sachar, D. Discriminant analysis application to understand the usage of digital channels while buying a car. *South Asian J. Mark.* **2021**, *2*, 51–71. [CrossRef]
6. Pinto, P. Decoding customer concerns about embracing electric cars in India: Analysis of audience sentiments on YouTube autovlogs. *Shanlax Int. J. Arts Sci. Humanit.* **2023**, *11*. [CrossRef]
7. Usmani, A.; Alsamhi, S.H.; Khan, M.J.; Breslin, J.; Curry, E. MuSe-CarASTE: A comprehensive dataset for aspect sentiment triplet extraction in automotive review videos. *Expert Syst. Appl.* **2025**, *262*, 125695. [CrossRef]
8. Hua, Y.C.; Denny, P.; Wicker, J.; Taskova, K. A systematic review of aspect-based sentiment analysis: Domains, methods, and trends. *Artif. Intell. Rev.* **2024**, *57*, 296–330. [CrossRef]
9. Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Chi, E.; Le, Q.; Zhou, D. Chain-of-thought prompting elicits reasoning in large language models. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 24824–24837. [CrossRef]
10. Wang, B.; Zhang, Z.; Li, X.; Jiang, L.; Chen, S.; Wang, B.; Zhou, J. Grammar prompting for domain-specific language generation with large language models. *Adv. Neural Inf. Process. Syst.* **2024**, *36*, 65030–65055. [CrossRef]
11. Serrano-Guerrero, M.; Bani-Doumi, F.P.; Romero, F.P.; Olivas, J.A. A 2-tuple fuzzy linguistic model for recommending healthcare services grounded on aspect-based sentiment analysis. *Expert Syst. Appl.* **2024**, *238*, 122340. [CrossRef]
12. Carrasco-Aguilar, A.; Galán, J.J.; Carrasco, R.A. Obamacare: A bibliometric perspective. *Front. Public Health* **2022**, *10*, 979064. [CrossRef] [PubMed]
13. Cobo, M.J.; López-Herrera, A.G.; Herrera-Viedma, E.; Herrera, F. SciMAT: A new science mapping analysis software tool. *J. Am. Soc. Inf. Sci. Technol.* **2012**, *63*, 1609–1630. [CrossRef]
14. Xu, H.; Zhao, P.; Li, P.; Zhang, Y.; Chen, H. Dual-enhanced generative model with graph attention network and contrastive learning for aspect sentiment triplet extraction. *Knowl. Based Syst.* **2024**, *301*, 112342. [CrossRef]
15. Li, Q.; Wen, W.; Qin, J. Improving span-based aspect sentiment triplet extraction with part-of-speech filtering and contrastive learning. *Neural Netw.* **2024**, *177*, 241–258. [CrossRef]
16. Hu, Z.; Wang, Z.; Wang, Y.; Tan, A. Aspect sentiment triplet extraction incorporating syntactic constituency parsing tree and commonsense knowledge graph. *Cogn. Comput.* **2023**, *15*, 337–347. [CrossRef]
17. Jin, B.; Zhang, Z.; Wang, Y.; Chen, X.; Zhao, J. Span-based dependency-enhanced graph convolutional network for aspect sentiment triplet extraction. *Neurocomputing* **2024**, *564*, 126966. [CrossRef]
18. Yuan, L.; Wang, J.; Yu, L.; Zhang, X. Encoding syntactic information into transformers for aspect-based sentiment triplet extraction. *IEEE Trans. Affect. Comput.* **2024**, *15*, 722–735. [CrossRef]
19. Liu, H.; Zhang, C.; Xia, R.; Cambria, E. Exploiting duality in aspect sentiment triplet extraction with sequential prompting. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 6111–6123. [CrossRef]
20. Xu, Q.; Chen, J.; Zhang, H.; Wang, H.; Li, H.; Zhang, J.; Zhang, M. DTDA: Dual-channel triple-to-quintuple data augmentation for comparative opinion quintuple extraction. *Knowl. Based Syst.* **2025**, *307*, 112734. [CrossRef]
21. Li, Y.; Zhang, S.; Li, H.; He, Z.; Tang, J. A span-sharing joint extraction framework for harvesting aspect sentiment triplets. *Knowl. Based Syst.* **2022**, *242*, 108366. [CrossRef]

22. Li, Y.C.; Wang, F.; Zhong, S. A more fine-grained aspect-sentiment-opinion triplet extraction task. *Mathematics* **2023**, *11*, 3165. [CrossRef]
23. Ballas, V.; Tsimpourlas, E.; Kenteris, M.; Sampson, D. Automating mobile app review user feedback with aspect-based sentiment analysis. *Hum. Cent. Des. Oper. Eval. Mob. Commun.* **2024**, *14738*, 179–193. [CrossRef]
24. Tonis, R.B.M. Automating scientific paper screening with Backus-Naur Form (BNF) grammars. *Didact. Danub.* **2024**, *4*, 46–57.
25. McKeeman, W.M. BNF programming, analyzing, and compiling. *Stud. Comput. Sci. Artif. Intell.* **1988**, *4*, 201–237. [CrossRef]
26. Zadeh, L.A. The concept of a linguistic variable and its application to approximate reasoning—I. *Inf. Sci.* **1975**, *8*, 199–249. [CrossRef]
27. Herrera, F.; Martínez, L. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Trans. Fuzzy Syst.* **2002**, *8*, 746–752. [CrossRef]
28. CleanTechnica. Tesla Sweeps Podium for Top-Selling Electric Vehicles in USA. Available online: https://cleantechnica.com/2024/12/09/tesla-sweeps-podium-for-top-selling-electric-vehicles-in-usa-charts/ (accessed on 29 December 2024).
29. Carwow. Car Reviews, Car News & Buying Advice. YouTube. Available online: https://www.youtube.com/user/Carwow (accessed on 29 December 2024).
30. Depox, J. YouTube Transcript API, Version 0.6.3. Available online: https://pypi.org/project/youtube-transcript-api/ (accessed on 29 December 2024).
31. Carwow. Tesla Model 3 in-Depth Review-See Why It's the Best Electric Car in the World! YouTube. Available online: https://www.youtube.com/watch?v=kbulCM90w8w (accessed on 29 December 2024).
32. Carwow. Tesla Model Y 2022 Review-The BEST Electric SUV? YouTube. Available online: https://www.youtube.com/watch?v=PZ8NPeYFPCY (accessed on 29 December 2024).
33. Carwow. Hyundai Ioniq 5 Review with 0–60mph Test! YouTube. Available online: https://www.youtube.com/watch?v=3mgh_1-1sTU (accessed on 29 December 2024).