Quot Homines, Tot Sententiae? Estimating Number of Different Opinions in Product Reviews

Nadezhda Chechneva Saint Petersburg State University Saint Petersburg, Russia chechnevanadegda@mail.ru

Abstract- This paper explores the diversity of people's opinions. We have chosen as the research material a corpus of customer reviews on robotic vacuum cleaners in the Russian language. The study aims to identify unique perspectives and recurring patterns in user feedback, examining how individuals express their satisfaction or dissatisfaction with the product. In particular, we estimate the growth rate of the number of different opinions with an increase of the corpus size. By employing both qualitative and quantitative methods, the research explores the linguistic strategies used by users to convey their experiences and emotions. We utilize a large language model (LLM) as a tool for extracting and analyzing opinions from product reviews.

I. INTRODUCTION

"Quot homines, tot sententiae" ("as many men, as many opinions") – this is well-known Latin phrase, attributed to the Roman playwright Terence, it captures the diversity of human thought and perception. The phrase suggests that every individual is unique in their perspective, influenced by their personal experiences, beliefs, and expectations. In this research, we aim to explore how this centuries-old adage holds up in a modern context, specifically focusing on the diversity of opinions that emerge in reviews of a single product. By analyzing user feedback on robotic vacuum cleaners, we seek to understand the extent to which opinions on a relatively simple and functional product can differ.

Robotic vacuum cleaners were selected as the subject of this study because of their limited functionality and clear, practical purpose. They are designed to perform a straightforward task – cleaning floors – and typically have few complex features compared to other advanced consumer technologies. This simplicity makes robotic vacuums an ideal product for investigating the range of user opinions, as it is interesting to observe how many unique and varied opinions can arise in relation to such a utilitarian object. How do users articulate their satisfaction or dissatisfaction with a product that offers a limited set of functions? Are opinions largely uniform, or do we observe a wide diversity even in this narrow domain? These questions form the basis of our inquiry.

Reviews are traditionally viewed from a more practical perspective: researchers are primarily interested in how they help consumers make informed decisions when selecting products and how they provide manufacturers with valuable feedback to improve their offerings [1, 2, 3, 4]. However, in this study, we focus on a deeper and less explored dimension– the nature and diversity of opinions themselves. Rather than merely considering reviews as functional tools for decisionmaking or product enhancement, we are interested in the underlying linguistic and cognitive aspects that reveal how individual experiences and perceptions shape unique viewpoints. By examining the diversity of these perspectives, we aim to uncover patterns in how users express satisfaction or dissatisfaction, and how even seemingly similar experiences can give rise to a wide range of opinions.

To achieve this, we utilize a large language model (LLM) as a tool for extracting and analyzing user opinions. Recent advances in natural language processing (NLP), particularly through LLMs, have opened new possibilities for understanding and interpreting large volumes of text data.

Studies have shown that LLMs can effectively classify sentiments in reviews and categorize feedback by product features. For instance, the paper [5] describes the application of LLMs such as GPT-3.5, GPT-4, and Llama 2 in the context of sentiment analysis, a key task in marketing research used to interpret consumer emotions and opinions. The study benchmarks these LLMs against traditional transfer learning models and finds that LLMs, even in zero-shot scenarios, can match or surpass these established methods in terms of classification accuracy.

Similarly, the paper [6] takes a detailed look at how well large language models (LLMs) can handle different types of sentiment analysis tasks. These tasks range from basic sentiment classification (like determining whether a review is positive or negative) to more advanced tasks, such as analyzing specific aspects of a product's sentiment (aspect-based sentiment analysis) and understanding complex subjective opinions. The study tests LLMs on 13 different tasks using 26 datasets, and compares their performance to smaller language models (SLMs) that are trained for specific domains.

LLMs have been shown to effectively identify patterns, extract structured information, and classify sentiment in a wide range of textual sources, including product reviews. By leveraging these capabilities, we aim to investigate the uniqueness and variation in user feedback, focusing on the linguistic and emotional aspects of these reviews.

Thus, the research is aimed at identifying the extent of opinion variability in reviews and attempting to classify these opinions in terms of their uniqueness. This approach allows us to view reviews not only as a tool for evaluating product quality but also as a valuable source of data on the depth and complexity of user perceptions.

II. SEMANTIC NATURE OF OPINIONS

The object of this research is the linguistic means of expressing opinions in consumer reviews. An opinion, as a semantic concept, implies the expression of one's stance or position regarding something. In the explanatory dictionary of the Russian language, the lexical meaning of the word "opinion" is described as follows: "a judgment expressing an evaluation of someone or something, an attitude toward someone or something, a view on someone or something" [7]. However, such a brief definition does not capture the multifaceted nature of this concept, which encompasses a wide range of units in language. To better understand the essence of opinions, we will briefly review the history of opinion studies and their connection to other concepts.

The question of what constitutes an opinion falls within the realm of epistemology, as it is a mental category. According to philosophical tradition, opinions are typically contrasted with another mental category–knowledge. The dichotomy of "knowledge-opinion" has been extensively studied in linguistics as well.

Initially, in descriptions of the "knowledge-opinion" dichotomy, the central concept was "knowledge": researchers focused on its defining characteristics, such as verifiability, comprehensibility, and truthfulness, while "opinion" was understood as that which does not meet the criteria of knowledge [8]. Over time, the trend of opposing knowledge and opinion persisted, with attempts to identify more specific differences.

Y.D. Apresyan [9] explains the primary distinction between knowledge and opinion, which lies in their relation to the truth of the statements in which they are used. Knowledge belongs to the class of factive predicates and retains truth even when negated in the subordinate clause. For example, in the sentences "He knew that his friends had betrayed him" and "He didn't know that his friends had betrayed him," the fact of the friends' betrayal is reported equally. In contrast, opinions belong to the class of putative predicates and do not necessarily imply truth.

Apresyan also identifies several semantic differences between opinion and knowledge that arise from the main distinction:

- Knowledge is singular and unchanging, while opinions are subject to change and allow for different perspectives;
- Knowledge is depersonalized, while opinions are personalized and have a specific holder;
- Knowledge is derived from an external source, whereas opinions are formed by an act of human will;
- Knowledge is stored in memory, while opinions reside in the mind;
- Knowledge flows from the external world to the subject, while opinions tend to expand outward into the external world;
- The value of knowledge is determined by its source, while the value of opinions depends on the intellectual or social status of their holders.

At the same time, some scholars do not draw a clear line between opinion and knowledge, allowing for overlap between these categories. N. D. Arutyunova writes that "the verification of a judgment transforms an opinion into knowledge, filling the gap that separates man from the world, for the world is knowable." Thus, a judgment "alienates itself from the individual and becomes objective truth" [10]. M. A. Dmitrovskaya in [8] also notes that many types of opinions tend to stabilize and transition into the realm of "personal" knowledge.

Beyond the comparison of opinions and knowledge, the relationship between opinions and evaluative judgments is also important. N. D. Arutyunova metaphorically emphasizes the value-laden nature of opinions: "The mode of belief (opinion) forms that shaky bridge which man can throw from the world of values (the world of the desirable and the ought-to-be) into the physical world" [10]. She identifies the following key properties of evaluations and opinions:

- Evaluation has no cause, while an opinion may have one; moreover, the motive behind an evaluation can be understood as the cause of an opinion. For example, Arutyunova notes that while we cannot ask, "Why are roses beautiful?" we can ask, "Why do you think roses are beautiful?";
- Evaluation may be motivated by an opinion but cannot be verified.

G.F. Ivanova [11] also discusses the relationship between opinions and evaluative judgments. She points out that both opinions and evaluations are judgments in which the speaker expresses their view on the state of affairs in the world. Ivanova asserts that evaluations clearly fall within the realm of opinions, as they reflect an individual's life experience. When formulating an evaluation, the speaker does not describe the possibility of a state of affairs but rather their viewpoint on the perceived, objectively existing objects, phenomena, and events.

A less obvious but significant aspect of opinion research is its connection to emotions. This question is debated and interpreted differently by various scholars. According to E.M. Volf [12], some cognitive researchers suggest that opinions may precede emotions and are a necessary part of emotional formation. Additionally, certain opinions may be identical to emotions, while emotions, in turn, may give rise to specific opinions. Research indicates that evaluative and factive opinions serve as logical presuppositions of emotions, each emotion having a typical set of such opinions. For instance, the statement "I am angry at my sister" conveys not only the speaker's emotional state but also a negative judgment about the sister's behavior.

In conclusion, the boundaries between opinions, emotions, and evaluations are often blurred and intertwined. In this study, we focus on opinions expressed in consumer reviews. For this type of text, as shown in the next section, the expression of the author's evaluative stance towards the product, including emotionally charged statements, plays an important role. Therefore, we deliberately broaden our scope to include not only explicit evaluations but also statements with evaluative connotations. Thus, the research is aimed at identifying the extent of opinion variability in reviews and attempting to classify these opinions in terms of their uniqueness.

III. PROBLEM DEFINITION

Within this study, we define an opinion as a combination of an aspect and a tonal marker. The aspect refers to a specific part or characteristic of the product, or the product itself, to which the opinion is directed. The tonal marker is a word or phrase that indicates the emotional or evaluative tone of the sentence or statement.

It is evident that the same opinion can be expressed using different words. This raises the question of what constitutes a unique opinion versus what should be considered contextually synonymous. To address this, we have developed a set of rules to determine when two opinions should be considered contextually synonymous:

1) They refer to the same aspect or object.

2) They share the same evaluative polarity (positive, negative, or neutral).

3) Both convey approximately the same amount of new information about the object.

For example, the phrases "vacuuming well" and "does an excellent job with its primary function" are contextually synonymous because they refer to the same aspect–vacuuming performance–have the same positive polarity, and convey roughly the same level of new information. However, "vacuuming pet hair effectively" would not be considered synonymous, as it introduces additional information (specifically, its ability to vacuum pet hair), even though it relates to the same aspect and shares the same positive polarity.

This approach allows us to differentiate between opinions that are merely different in wording but identical in meaning, and those that provide distinct insights into the product's features or performance. By classifying opinions in this way, we can more accurately assess the range and depth of user perceptions and their uniqueness.

IV. EXTRACTING CONSUMER OPINIONS FROM PRODUCT REVIEWS USING A LLM

In this section, we describe the process of extracting consumer opinions from product reviews using a large language model (LLM). The main goal was to prompt the model to not only extract the opinions but also to identify contextually synonymous opinions based on certain criteria. The model was instructed to unify the expressions of contextually similar opinions, ensuring consistency in the analysis. We used Open AI GPT-40 model [13].

A. Prompt for the LLM

We used a carefully crafted prompt to guide the LLM in extracting opinions. The task required the model to recognize when different phrasings represented the same underlying opinion by focusing on three core criteria:

- Same Aspect or Object: Both opinions should refer to the same product feature or aspect.
- **Same Sentiment Polarity**: Both opinions should share the same sentiment (positive, negative, or neutral).

• Similar Information Content: Both opinions should provide approximately the same amount of new information about the aspect.

Since the corpus of the reviews is in Russian we interacted with the LLM in Russian. The following is an English translation of the exact prompt used to instruct the LLM:

Prompt: "You are an assistant helping to extract consumer opinions from reviews. Your task is to learn how to identify contextually synonymous opinions. Two opinions are synonymous if:

- They refer to the same aspect/object.
- They share the same sentiment polarity (positive, negative, neutral).
- They contain approximately the same amount of new information about the object. For example, 'vacuums well' and 'performs its main function excellently' are contextually synonymous, as they refer to the same aspect (vacuuming), share the same positive polarity, and offer a similar amount of information. On the other hand, 'vacuums pet hair well' is not synonymous because it adds additional information (specifically regarding pet hair). Extract the segment containing the opinion, as well as the opinion itself. Reformulate opinions so that contextually synonymous ones are expressed identically. If a sentence contains multiple opinions, extract them separately."

Moreover, to ensure clarity and precision, the model was provided with concrete examples of how opinions should be extracted and normalized.

B. Output Format

The model was instructed to provide extracted opinions in a consistent and structured format, using the following notation:

- **t:** The segment of the review text containing the opinion.
- **o:** The reformulated opinion.
- **p:** The polarity of the opinion ("+" for positive, "-" for negative, "+ -" for neutral).

This structured approach enabled the consistent extraction of opinions across various reviews, regardless of phrasing differences.

C. Handling Multiple Opinions

In cases where a sentence contained multiple opinions about different aspects, the LLM was instructed to extract each opinion separately. For example:

"The vacuum is fast and quiet, but the app is slow."

The output would be:

t: "The vacuum is fast", o: "fast vacuuming", p: "+"

t: "and quiet", o: "quiet operation", p: "+"

t: "but the app is slow", o: "slow app", p: "-"

This method allowed for a nuanced and comprehensive extraction of opinions, ensuring that every distinct opinion was captured and classified correctly.

D. Results

We analyzed 1,000 reviews of robot vacuum cleaners, and the model extracted 2,661 fragments containing opinions, from which 1,355 unique opinions were identified. The analysis revealed a significant number of repeated opinions, meaning different phrases conveyed similar or identical meanings. The chart in Fig.1 illustrates the distribution of the total reviews, fragments extracted, and unique opinions. This finding highlights the prevalence of paraphrased expressions that, while different in wording, share the same underlying sentiment and evaluation of product aspects.

The extracted opinions demonstrate a notable degree of repetition, where different formulations express the same underlying sentiment or evaluation. This indicates that despite varied linguistic expressions, the core opinion remains the same. For example:

- "качественная уборка" ("high-quality cleaning") and "хорошо убирает" ("cleans well") – both refer to the overall cleaning performance of the robot vacuum.
- "хорошо преодолевает препятствия" ("good at overcoming obstacles") and "преодолевает препятствия на ура" ("handles obstacles like a pro")
 both opinions refer to the robot's capability to navigate obstacles.
- "наличие функции влажной уборки" ("presence of the mopping function") and "функция влажной уборки" ("mopping function") – both point to the availability of the mopping feature, even though expressed differently.

Such repetitions emphasize the need for careful filtering and grouping of opinions to avoid overestimating the diversity of sentiment. In our analysis, these contextually synonymous opinions were identified and grouped, ensuring that opinions with the same meaning but different wording were not counted multiple times.

V. CLUSTERING SIMILAR OPINIONS

We approached the next stage of analysis with the goal of further reducing the number of opinions identified in the previous step. After the LLM had extracted 1,300 unique opinions, we aimed to refine this dataset by having the model identify groups of opinions that, while worded differently, conveyed the same meaning. For this, all of the unique opinions were fed into the model, and it was tasked with finding synonymous opinions, expressed in different ways. This process was repeated over five iterations, each time asking the LLM to perform the same task of identifying synonymous opinions. As a result, the original list of opinions was reduced from 1,355 to 413.



Fig. 1. Opinion grouping stages and unique opinions count

The diagrams in Fig.2 and Fig.3 illustrate examples of opinion clustering.





Fig. 2 demonstrates the consolidation of multiple user reviews into the final category of "Mediocre quality of wet cleaning".



Fig. 3. Clustering of user opinions regarding the effectiveness of pet hair cleaning

Fig. 3 illustrates the grouping of various user reviews into the overarching category of "Quality pet hair cleaning".

VI. SATURATION POINT IN OPINION GROWTH

In this part of the study, we aimed to understand how the number of unique opinions grows as the corpus of reviews increases. The graph shows the relationship between the number of unique reviews and the corresponding number of unique opinions extracted. Our primary objective was to identify the point of saturation, where adding more reviews no longer significantly increases the number of unique opinions.



Fig. 4. Growth of unique opinions with increasing unique reviews

From the graph, we observe an initial steep rise in unique opinions as the number of reviews increases. This suggests that early in the process, each new review introduces a significant number of novel perspectives. However, as the review count approaches 500, the rate of growth in unique opinions begins to slow down, indicating that many of the commonly held opinions have already been expressed. Beyond 600 reviews, the curve flattens further, revealing a point of diminishing returns. At this stage, additional reviews contribute fewer new opinions, suggesting that the saturation point is near.

In conclusion, the graph shows that while the number of unique opinions grows with increasing reviews, there is a clear point where this growth slows, signaling a saturation of opinions around 400 unique opinions. This insight can help streamline future analysis by focusing on a smaller yet representative sample of reviews, knowing that additional reviews may add little new information.

VII. OPINION CLASSES

At the final stage of our research, we aim to validate whether the six thematic aspect classes proposed in our previous work [14] are suitable for categorizing unique opinions extracted from reviews. In that article, we introduced six thematic aspect classes, which are universal for a broad range of products, including digital and household appliances. These categories were derived based on the existing conceptual representation of the technical devices. Now, let's briefly describe each of the six thematic classes:

• Functional characteristics – these are characteristics related to the product's intended purpose. This includes descriptions of the functions performed by the device, how effectively they are executed, how well they align with the requirements for fulfilling the product's purpose, and the ability to select, configure, or combine functions required for the target purpose.

- Operational characteristics these indicators reflect the possible and actual use of the product under existing operating conditions. This includes ease of use, reliability, durability, maintainability, safety, technical perfection, degree of automation, as well as device-specific parameters that significantly depend on the device's type and purpose, such as power, performance, and energy consumption.
- Construction a description of the product's components, structure, dimensions, weight, materials, and assembly.
- Aesthetic characteristics descriptions of the product's appearance, design, color, shape, and decorative elements.
- Price characteristics related to price and the price-toquality ratio.
- General characteristics general evaluations of the product as a whole, as well as features that do not fall into the other categories.

To further evaluate the applicability of the six thematic aspect classes, we employed a large language model (LLM) to automatically assign each unique opinion to one of the six predefined categories. The resulting distribution of opinions across the six thematic classes is displayed in the graph below. This visualization provides insight into how user feedback is categorized, highlighting trends in the prominence of certain aspects over others.



Fig. 5. Distribution of opinions by thematic classes

Functional characteristics receive the highest number of positive opinions, with 949 positive mentions and 105 negative mentions. This suggests users are generally satisfied with the functionality of the product.

Operational characteristics have the most negative feedback among all aspects, with 414 negative opinions compared to 357 positive ones. This indicates significant user concerns about the product's performance and usability.

Construction has more negative feedback (41) than positive (31), which may point to issues with build quality or design.

Aesthetic characteristics are viewed mostly positively, with 21 positive opinions and no negative feedback, indicating satisfaction with the product's design.

Price shows more positive feedback (94) than negative (5), suggesting good value for money in the eyes of users.

General characteristics also show a strong majority of positive opinions (363 positive vs. 16 negative), indicating overall user satisfaction with the product.

The graphs in Fig.6 and Fig.7 illustrate the distribution of the top 20 most frequent opinions in the largest and most frequently discussed by users thematic aspect classes: *Functional Characteristics* and *Operational Characteristics*.



Fig. 6. Top 20 Opinions in Functional Characteristics

The graph in Fig.6 highlights that the majority of user feedback is focused on the product's cleaning performance. The most common opinions include "High-quality cleaning", "High-quality wet cleaning", and "High-quality pet hair cleaning", demonstrating that effective cleaning is the most critical functional expectation for users. Other significant mentions include navigation, suction power, and the device's ability to maintain daily cleanliness.

There are also some negative opinions within the top 20, such as "ineffective wet cleaning" and "poor carpet cleaning", which suggest variability in user experiences with the product's cleaning capabilities.

For the category *Operational Characteristics*, the graph indicates that ease of use and practical performance are the most important aspects for users. The most frequent opinion is "convenient to use," followed by feedback on "quiet operation," "overcoming obstacles," and "battery charge" issues. Users also provide a mix of positive and negative feedback regarding operational features, with concerns like "inconvenient mobile app" and "Russian localization issues" standing out.



Fig. 7. Top 20 Opinions in Operational Characteristics

VIII. CONCLUSION

This study has provided a comprehensive analysis of the diversity of opinions expressed in user reviews for robotic vacuum cleaners. Overall, this research has revealed that even for a relatively simple product like a robotic vacuum cleaner, the diversity of user opinions is remarkably broad. By employing a large language model (LLM) for the extraction and classification of opinions, we were able to investigate not only the sentiments associated with specific product features but also the variability and uniqueness of user feedback.

The saturation analysis further demonstrated that the growth of unique opinions slows significantly after a certain number of reviews, indicating that adding more reviews beyond this point contributes little new information. This finding is crucial for optimizing future analyses by focusing on a representative subset of reviews. We hypothesize that in more sophisticated or feature-rich devices, users may continue to express new opinions even after analyzing a large corpus of reviews. This means that the saturation point, where new unique opinions plateau, could occur much later for these types of products. This variation in saturation thresholds for different products could serve as a subject of further study, as it would offer deeper insights into how consumers interact with and evaluate a broader and more complex array of product features. Understanding these dynamics could help optimize review analysis strategies for products with varying levels of complexity.

The thematic aspect classes introduced in our previous work have proven effective in categorizing user feedback.

In conclusion, this research contributes to a deeper understanding of how customers articulate their experiences with relatively simple products, such as robotic vacuum cleaners. Also the study demonstrates the potential of LLMs in exploring the diversity of people's opinions. Future work could expand this approach to more complex products and touch additional dimensions of user feedback.

REFERENCES

- S. Das, M. Chen "Yahoo! for Amazon: Extracting market sentiment from stock message boards", in: Proceedings of the Asia Pacific finance association annual conference (APFA), T. 35, 2001, pp. 1375-1388.
- [2] K. Dave, S. Lawrence, D.M. Pennock "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews", in: Proceedings of the 12th international conference on World Wide Web, 2003, pp. 519-528.
- [3] T. Nasukawa, J.Yi "Sentiment analysis: Capturing favorability using natural language processing", in: Proceedings of the 2nd international conference on Knowledge capture, 2003, pp. 70-77.
- [4] R.M. Tong "An operational system for detecting and tracking opinions in on-line discussion", in: Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification, T. 1, №. 6, 2001, pp.1-6.
- [5] J.O. Krugmann, J. Hartmann, "Sentiment Analysis in the Age of Generative AI", Customer Needs and Solutions, vol.11, 2024.
- [6] W.Zhang, Y.Deng, B.Liu, S.Pan, and L.Bin, "Sentiment Analysis in the Era of Large Language Models: A Reality Check". In Findings of the Association for Computational Linguistics: NAACL 2024, June 2024, pp. 3881–3906.
- [7] S.A. Kuznetsov, Noveishii bol'shoi tolkovyi slovar' russkogo yazyka [The Newest Large Explanatory Dictionary of the Russian Language], Moscow: Ripol klassik [Ripol Classic], 2008.
- [8] M.A. Dmitrovskaya, "Znanije i mnenije: obraz mira, obraz cheloveka" [Knowledge and Opinion: The Image of the World, the

Image of a Person], Logicheskii analiz yazyka. Issue 1. Znanije i mnenije [Logical Analysis of Language. Issue 1. Knowledge and Opinion], Moscow: Nauka [Science], 1988, pp. 7-17.

- [9] Y.D. Apresyan, "Sistemaobrazuyushchie smysly 'znat' i 'schitat' v russkom yazyke" [System-Forming Meanings of 'Know' and 'Consider' in the Russian Language], Russkii yazyk v nauchnom osveshchenii [Russian Language in Scientific Light], vol. 1, 2001, pp. 5-26.
- [10] N.D. Arutyunova, Tipy yazykovykh znachenii. Otsenka. Sobytie. Fakt [Types of Linguistic Meanings: Evaluation, Event, Fact], Moscow: Nauka [Science], 1988.
- [11] G.F. Ivanova, "O mneniyakh i otsenkakh" [On Opinions and Evaluations], Izvestiya Rossiiskogo gosudarstvennogo pedagogicheskogo universiteta im. A. I. Gertsena [Izvestiya of the Russian State Pedagogical University named after A.I. Herzen], vol. 8, no. 41, 2007, pp. 25-31.
- [12] E.M. Volf. Funktsional'naya semantika otsenki [Functional Semantics of Evaluation], 2nd ed., revised and expanded, Moscow: Editorial URSS, 2002.
- [13] J. Achiam et al. "Gpt-4 technical report", arXiv preprint arXiv:2303.08774, 2023.
- [14] N.S. Chechneva, "Ob odnom podkhode k avtomaticheskoi summarizatsii potrebitel'skikh otzyvov" [On an Approach to Automatic Summarization of Consumer Reviews], Vestnik NSU. Seriya: Lingvistika i mezhkul'turnaya kommunikatsiya [Bulletin of NSU. Series: Linguistics and Intercultural Communication], 2022, pp. 90-106.