Demographic-Based Public Perception Analysis of Electric Vehicles on Online Social Networks

Tavishi Priyam †, Tao Ruan *†,‡ and Qin Lv *

Department of Computer Science, University of Colorado Boulder, Boulder, CO 80309, USA;
tavishi.priyam@colorado.edu
* Correspondence: tao.ruan@colorado.edu (T.R.); qin.lv@colorado.edu (Q.L.)
† These authors contributed equally to this work.

Abstract: Electric vehicles have gained significant popularity in the market, with sales increasing yearly. The introduction of new policies and reforms aimed at promoting environmental sustainability, coupled with the release of more advanced electric vehicles with higher driving ranges and technical specifications, has encouraged more people to consider switching to electric vehicles. However, there is still a lack of understanding of public perception and the factors influencing the decision to switch to electric vehicles, especially among people from different demographic groups. In this study, we leverage machine learning techniques to analyze public opinion about electric vehicles across different demographic groups on two online social networks (OSNs), namely Reddit and Twitter. Our analyses provide valuable insights into how users on these platforms perceive electric vehicles and the factors that influence their perception. This information can be used to inform market strategies and future policies aimed at promoting the adoption of electric vehicles.

Keywords: social networks; demographics; electric vehicles; public perception; case study

1. Introduction

The transportation sector is responsible for approximately 20% of global carbon emissions, and in the United States, it contributes to 29% of carbon emissions [1,2]. Consequently, the rapid adoption of electric vehicles (EVs) is regarded as a crucial element of decarbonization efforts by various climate change mitigation plans [3]. For instance, the Biden administration aims to achieve EV sales of between 40% and 50% by 2030 [4,5], while the United Nations has set the objective of achieving carbon neutrality by 2050 [6].

In recent years, annual global EV sales have been firmly on the upswing, and the market share of electric vehicles (EVs) reached a record high in 2021 with nearly 9% of the total market share, according to a report by the International Energy Agency [7]. This growth in EV sales is attributed to various factors, such as energy efficiency, cost savings, policy, and convenience.

Despite this impressive growth, there is still a long way to go before EVs achieve mass-market adoption. One issue is the limited understanding of how different demographic groups perceive and approach the adoption of EVs. Prior research suggests that socio-demographic attributes, including age, gender, and political orientation, play a crucial role in potential EV adopters [8,9]. However, traditional survey-based methods employed in these studies have several limitations, such as small sample sizes, limited timeframes, and the “Hawthorne effect” [10,11]. In this work, our specific objective is to comprehensively analyze the perception of electric vehicles across varied demographic groups, focusing on how factors like age, gender, and political orientation influence attitudes and discussions related to EVs.

To achieve this, we have delineated our study to include an in-depth investigation of three primary demographic variables: age, gender, and political orientation. We leverage the vast amount of data available on two different online social networks (OSNs), Twitter and...
and Reddit, to study how different demographic groups perceive EVs and the factors that affect their inclination toward EVs. This approach allows us to not only gain a broad understanding of the overall public perception but also to understand the nuanced differences across these demographic segments.

Our research focuses on individuals belonging to different age and gender groups. We also investigate online communities with different political leanings to determine whether there are any differences in how these groups approach discussions about EVs. Through this approach, we aim to reveal detailed insights into the specific aspects of public perception of EVs, including sentiment, prevalent themes, and focus areas of discussion, across these demographic groups. This will contribute to a more comprehensive and nuanced understanding of the factors influencing the adoption of EVs among diverse populations. Our main contributions are multi-faceted:

- Our work represents the first study that utilizes large-scale online social network (OSN) datasets to analyze demographic-based public perception of EVs with the help of the latest machine learning inference approach.
- We conducted a gender-based analysis of posts from Twitter users by classifying users using the M3 model, which revealed a higher share of voice for male compared to female users in EV-related discussions on OSNs. Nevertheless, users from both genders hold similar sentiments toward EVs.
- We conducted an age-based analysis across Twitter users, revealing that nearly 50% of the voices belonged to individuals over 40 years of age. The topic modeling analysis indicates that younger users focus more on the environmental aspect of EVs, while older users focus more on the administrative and infrastructural aspects.
- Our political analysis between left-leaning and right-leaning OSN users demonstrates that left-leaning discussions are centered around EVs’ political and environmental aspects. In contrast, right-leaning discussions are more focused on the financial aspects.
- Our analysis complements the previous large-scale Fuel Institute Consumer Survey [12]. We find that our gender and political analysis is consistent with the survey. Additionally, the results from our political analysis outlined in this paper indicate that both left- and right-leaning OSN users hold a positive opinion of EVs, with the left-leaning online community having a higher degree of positive sentiment.
- Our case study between the online discussions in the Colorado and Utah Reddit communities indicates that political affiliation alone is insufficient to determine public opinion of EVs. The location is an important feature to factor into analyses to understand user opinions.

2. Literature Review

2.1. Social–Demographic Factors in EV Perception

Several previous studies have explored the role of socio-demographic attributes in electric vehicle adoption. For example, Chen et al. investigated socio-demographic, behavioral, economic, and technical factors associated with electric vehicles based on a survey administered to 4,885 participants across 17 cities in the Nordic countries [8]. Comprehensive reviews of studies examining demographic factors in EV adoption have also been conducted [13,14]. Researchers have pointed out that social–demographic attributes, such as age and gender, are essential factors in potential adopters and are more critical than other attributes [8]. However, most of these studies rely on questionnaires, interviews, and customer surveys. These survey-based datasets are often limited in sample size and collected in a specific area within a relatively short time frame, such as a few weeks or months.

2.2. OSNs for Public Perception

Online social networks (OSNs) have become an integral part of people’s daily life and a significant facet of the public sphere. As primary hubs for public expression and interaction, OSNs offer a valuable source for researchers to study public perception in various domains,
such as political science [15,16], economics [17], and crisis informatics [18,19]. The concept of “text as data” has also gained importance in computational social science [20,21]. In the domain of transportation, OSNs have garnered increasing attention [22–25]. Zayet et al. conducted a review of transportation-related research based on social media analysis from 2008 to 2018 [26]. However, the utilization of OSNs to investigate the public perception of electric vehicles (EVs) is relatively scarce [27,28], partly because EVs have become more prevalent in recent years.

This study aims to bridge this research gap by utilizing spontaneous discussions on OSNs over a prolonged period to analyze the public perception of EVs.

2.3. Demographic Attributes on OSNs

Due to privacy concerns, OSNs typically do not provide demographic attributes such as age and gender in their datasets. Nonetheless, some studies have developed machine learning algorithms to identify user demographics. Brandt et al. investigated social media user demographics and topic diversity in their research [29]. Their work focused on an international policy forum and its discussions. The authors’ analyses provided valuable insights into the importance of women and youth voices in governing restoration in Africa. They further explored the differences in engagement among demographic groups. Cesare et al. reviewed existing approaches to automated detection of demographic characteristics of social media users before 2017 [30]. This extensive literature on user identification detection methods emphasized three notions: supervised/semi-supervised learning, facial evaluation, and data matching. Deep learning methods have demonstrated their power in capturing semantic information for natural language processing (NLP) in recent years. Wang et al. proposed a deep learning method (M3 model) to infer demographics from multilingual, profile-only data with state-of-the-art accuracy. Their work achieved high accuracy in detecting Twitter users’ demographic attributes, including gender, age, and organization.

Some researchers also investigate the political leaning of users in OSN analysis. For instance, Folgado et al. conducted research to comprehend the political landscape of Spain [31]. They utilized data focused on Spain’s five main political parties and determined the political affiliation of a tweet using a fully-connected neural network (FCNN). They also performed sentiment analysis on tweets across different groups and studied the evolution of tweets and their sentiment in reaction to political events.

Our current work builds upon the existing research [27,32], which aims to understand EV-related discussions on OSNs. The authors extensively analyzed a vast amount of data to comprehend the EV-related topics discussed by OSN users, the sentiment they held toward EVs, and how opinions evolved from 2011 to 2020. This work is beneficial in comprehending the topics and sentiment of the discussion and demonstrating how these have evolved with the advancement of electrified transportation.

2.4. Computational Social Science Methodologies

The analysis of large digital data repositories, such as online social network (OSN) data, often employs methods under the computational social science umbrella. Hofman et al. outline these methodologies in their study, as referenced in [33]. In a similar vein, Zhang et al. explore common data-driven computational social science techniques for analyzing OSN data in their survey [34]. These sources indicate that topic modeling and sentiment analysis are integral to OSN analysis. Topic modeling uncovers the subjects that capture public interest, while sentiment analysis gauges public sentiment toward specific events or products.

In our study, we also employ key computational social science methods, including Latent Dirichlet Allocation (LDA) [35] for topic modeling and the VADER algorithm [36] for sentiment analysis, to explore OSN data. LDA, a pivotal technique in natural language processing, groups words with high co-occurrence likelihood in documents, revealing prevalent topics. We utilize the Gensim Python toolkit for LDA analysis. VADER, recognized for its effectiveness in social media sentiment analysis, assesses text sentiment...
by assigning polarity scores to words, offering insights into public feelings and emotions. These methodologies, grounded in substantial research, enable us to dissect public interest and sentiment on topics related to electric vehicles, leveraging large datasets from platforms like Twitter and Reddit.

3. Data

This study utilizes data extracted from previous research [27,32] to examine public perception of EVs, comprising data from Twitter and Reddit. Twitter, a micro-blogging platform allowing posts of up to 280 characters, serves as one of the data sources. In our study, the Twitter data were collected using the Twitter Academic API, with the time frame extending from January 2021 to February 2022. This approach follows the data collection methodology outlined in previous research [32], but focuses on a more recent dataset. For our analysis, we used approximately 300,000 posts from this time period.

Reddit, known for its news aggregation and user-created communities called subreddits, provides another dimension to our study. We sourced our Reddit data from Pushshift, an archive containing all Reddit data, including submissions and comments, from 2005 to 2022 [37]. Focusing on the last decade (January 2011–December 2020), we extracted electric vehicle-related content in our previous research [27], retaining 437,917 Reddit posts, which include 274,979 submissions and 3,162,938 comments.

Both datasets, encompassing extensive social media data from Twitter and Reddit, offer insights into the public perception of EVs in the last decade, enabling a comprehensive analysis across different demographic groups.

3.1. Demographic Groups

Given the nature of online social networks (OSNs) like Twitter and Reddit, which do not inherently provide detailed demographic information such as age and gender, our study employed the M3 (Multimodal Multilingual and Multi-attribute) system [38]. This deep-learning model, capable of extracting information about Twitter users’ age and gender across 32 languages, was a cornerstone of our demographic analysis. In this study, we utilized the Python implementation of this model using the m3inference package to extract users’ age and gender information. Our methodology is particularly significant in capturing the demographic diversity of social media discourse on electric vehicles (EVs). This method enabled us to create a nuanced understanding of the demographic landscape of OSN users discussing EVs on Twitter, which is critical to our study’s aim of analyzing public perception across different demographic lines.

Alongside this, we also incorporated an analysis of political leaning, a critical factor in understanding public perception of electric vehicles (EVs). This aspect of our demographic analysis was particularly enhanced by our use of Reddit data, leveraging the platform’s unique characteristics, such as user communities and discussion threads, which often align with specific political ideologies.

3.1.1. Age and Gender

As shown in Figure 1, the distribution of users across two genders is depicted in 300,000 posts. Due to the tool’s limitations, we could only classify users into two gender groups. This preliminary analysis suggests that male voices are more dominant on social media than female voices on the Twitter platform.

The distribution of users across different age groups in a total of 300,000 posts is presented in Figure 2. The figure indicates that users over the age of 40 have the largest share of voice. In contrast, users between the ages of 18 and 29 have the lowest representation in EV-related discussions on Twitter.
3.1.2. Political Leanings

To examine the differences in opinion among individuals with varying political affiliations, it is crucial to classify data into different political groups. Given the challenging nature of classifying textual data based on political leanings on social media platforms, we have opted to use the communities on Reddit. Reddit is a social media platform that features a plethora of community forums where users can engage in discussions.

Our study used ten years of Reddit data from 2011 to 2020, focusing on politically oriented subreddits. We divided the data into two broad demographic groups: left- and right-leaning. The right-leaning groups include subreddit communities such as r/Conservative, r/republicans, and r/republican, while the left-leaning groups include subreddit communities such as r/democrats, r/liberal, r/LibertarianLeft, r/LibertarianSocialism, and r/libertarian. Figure 3 presents the number of EV discussions observed in the left-leaning and right-leaning groups in our study.

![Gender Distribution](image1.png)

**Figure 1.** Gender distribution.

![Age Distribution](image2.png)

**Figure 2.** Age distribution.
4. Analysis

To examine the perceptions of various demographic groups toward electric vehicles (EVs), we conducted separate topic modeling and sentiment analyses.

4.1. Methodology

For topic modeling, we employed the Latent Dirichlet Allocation (LDA) model [35]. LDA is a widely used topic modeling technique in which each document contains multiple words and topics, and each topic contains several words. LDA aids in identifying the most heavily discussed topics in each document. In our study, we considered each demographic group as a separate document. We used the ldamodel function provided by the Python Gensim package [39], analyzing the top four topics for each document after conducting ten iterations of the LDA modeling process. We defined a broad category for each topic based on the keywords to improve our analysis. Before performing the LDA algorithm, we proceeded to clean and pre-process the data for our analysis. We minimized the number of features for analysis by eliminating any unnecessary columns in the dataset. Following the dimensionality reduction, we began pre-processing the input data, which entailed removing punctuation marks, emoticons, symbols, URLs, and stopwords. We utilized regex and various functions provided by the NLTK package in Python to execute these tasks [40].

For sentiment analysis, we utilized the Valence Aware Dictionary for Sentiment Reasoning (VADER), an online dictionary that is highly effective in classifying unlabeled textual data, particularly from social media [36]. VADER takes into account both the polarity and intensity of the text. It maps lexical features to sentiment scores, and the resulting sentiment score ranges from $-1$ to $1$, with $-1$ representing the most negative and $+1$ representing the most positive sentences. We employed the SentimentIntensityAnalyzer, which is part of the NLTK implementation of VADER in Python [41].

4.2. Gender Analysis

The most notable finding for gender analysis is that male users dominate Twitter discussions, according to the classification performed using the M3 system. Regarding topics of interest, as shown in Figure 4a,b, female users’ most dominant topic centers on EV charging infrastructure, as evidenced by Topic 1 and Topic 3, both of which contain keywords such as charger, station, battery, and gas. These keywords suggest discussions about charging convenience and infrastructure. Other significant topics for female users include pricing, trade, and brand names such as Tesla.

The topic models for male users also identify “charging infrastructure” as the most dominant topic across all data. However, other topics, such as vehicle technical specifications, road conditions, trade, and investment, have almost equal importance.
To supplement our gender analysis and topic modeling and better understand the sentiment and perspective of male and female EV users toward online discussions, we conducted a sentiment analysis on the two genders separately and plotted the results. Figure 4c,d indicate the sentiment distribution across the two genders. Although variations in sample size and preferred topics were observed between male and female users in online social networks, the sentiment distribution demonstrates a noteworthy similarity. Notably, 48% of the female data sample and 47% of the male data sample expressed a neutral opinion toward EVs. A total of 4% of female and male OSN users indicated an extremely negative opinion of EVs.

4.3. Age Analysis

Users were divided into four different age groups based on the classification output from the M3 system. We conducted a topic modeling analysis on each group to understand the focal points of EV discussions and how the pattern differs.

Figure 5a depicts the topic models generated for younger users, revealing a shared interest in environmental consciousness across both demographic groups. This is evidenced by the presence of keywords such as "planet," "save," "emissions," and "trees." Additionally, the brand Tesla emerged as a popular topic of discussion among younger users.

Figure 5b displays the discussion topics among middle-aged and older users, highlighting mentions of technical specifications and makes of vehicles, as well as some political
discussions involving President Biden. The most significant common theme across all age groups was the convenience of the charging infrastructure.

In addition to topic modeling, we performed sentiment analysis to enhance our examination of OSN users across various age groups. Figure 6 illustrates the graphs representing sentiment distribution for each age cohort. Comparable to the patterns identified in our gender analysis regarding sentiment distribution, we note a consistent trend across age groups, wherein the predominant sentiment leans toward neutrality. Specifically, users aged 40 and above exhibit the least polarization in their opinion toward EVs, with 49% expressing a neutral stance according to our sentiment distribution. Although the disparity with other age groups is marginal (<5%), this demographic stands out as exhibiting relatively uniform sentiments.

Figure 5. Age analysis.

Figure 6. Sentiment distribution for different age groups.
Taken together, our findings suggest that younger users exhibit greater concern for the environmental impact of EVs and their effect on the planet. In contrast, older age groups demonstrate a stronger focus on EVs’ political and technical aspects, including battery, driving range, charging infrastructure, convenience, and comparisons to other hybrid alternatives.

4.4. Political Leaning Analysis

To investigate the impact of political affiliation on the perception of EVs among OSN users, we categorized subreddit communities into two groups based on their political leanings: left-leaning and right-leaning. The selection of these subreddit communities was based on factors such as the name, bio, and participation of individuals in discussions.

Our preliminary analysis of various political subreddits suggests that there were more posts from right-leaning users than left-leaning users, as demonstrated in Figure 3.

The topic modeling analysis (Figure 7a,b) indicated that left-leaning users focused primarily on general political and policy discussions surrounding EVs, with tax benefits, policies, environment, emissions, and brand mentions of Tesla being some of the most heavily discussed topics. In contrast, right-leaning users discussed the financial aspect of EVs, including topics such as pay, buy, cost, high, and gas, as well as general political and policy discussions and mentions of President Trump and tax benefits.

We also conducted a sentiment analysis (Figure 7c,d) and found that left-leaning users had a predominantly positive sentiment toward EVs, whereas right-leaning users were less positive in their EV discussions.

Figure 7. Topic modeling and sentiment analysis results for political analysis.
5. Case Study: Colorado vs. Utah

To gain a more in-depth understanding of how political affiliation affects people’s views on electric vehicles, a case study was conducted on subreddit communities focused on people residing in Colorado and Utah. Since Colorado is predominantly Democratic (left-leaning) and Utah is mainly right-leaning [42,43], this difference was leveraged to understand users’ opinions toward EVs. Despite the political leaning of each state, both have been promoting the adoption of electric vehicles, which is evident in our analysis.

In the meanwhile, Colorado and Utah share several socio-economic characteristics that make them comparable and suitable for a comparative case study on electric vehicles (EVs), as shown in Table 1.

Table 1. Social–economic comparison between Colorado and Utah.

<table>
<thead>
<tr>
<th></th>
<th>Colorado</th>
<th>Utah</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population (2021)</td>
<td>5.821 M</td>
<td>3.338 M</td>
</tr>
<tr>
<td>Land Area</td>
<td>104,094 square miles</td>
<td>84,899 square miles</td>
</tr>
<tr>
<td>Median Household Income (2020)</td>
<td>USD 80,184</td>
<td>USD 79,133</td>
</tr>
<tr>
<td>Real GDP (2022)</td>
<td>USD 385.83 billion</td>
<td>USD 191.97 billion</td>
</tr>
</tbody>
</table>


- **Economic structure**: Both Colorado and Utah have diverse economies that extend beyond traditional industries [44,45]. They are known for having strong service sectors, including technology, healthcare, and tourism. This economic diversity can impact the adoption of new technologies such as EVs.
- **Population growth**: Both states have experienced significant population growth in recent years. This growth can influence transportation needs, infrastructure development, and energy consumption patterns, all of which are pertinent factors when studying EV adoption.
- **Urban–rural mix**: Both states feature a mix of urban and rural areas. Studying EV adoption in regions with diverse population densities provides insights into the challenges and opportunities associated with catering to different transportation needs.
- **Environmental consciousness**: Colorado and Utah are known for their residents’ awareness of environmental issues. The population’s environmental consciousness can play a crucial role in the acceptance and adoption of environmentally friendly technologies, including electric vehicles.
- **Terrain and climate**: The states share similarities in terms of geography and climate, with mountainous terrains and diverse weather conditions. These factors can impact the performance and range of electric vehicles, making a comparative study relevant for understanding regional challenges.
- **Public infrastructure**: Colorado and Utah have been investing in infrastructure development, including charging stations for electric vehicles. Studying the development and effectiveness of these charging networks can provide valuable insights into the role of infrastructure in fostering EV adoption.
- **Cultural and demographic factors**: Both states have a mix of demographic groups and cultural backgrounds. Understanding how different demographic factors influence EV adoption rates can contribute to a more comprehensive analysis.

The socio-economic similarities between Colorado and Utah, encompassing economic diversity, population growth, environmental consciousness, and geographic considerations, make them compelling candidates for a case study on electric vehicle adoption. Analyzing
these states together can provide valuable insights into the factors influencing EV adoption and the effectiveness of policies and initiatives in different regional contexts.

Based on our analysis of around 12,000 data points, there are more EV discussions in Colorado than in Utah. While the volume of political discussion is higher in Colorado, with almost three-quarters share of voice, it seems that the political discussion in Utah is much more focused on the administration and policies.

As we can observe in Figure 8a,b, the most dominant topic discussed in Colorado is the vehicle and charging infrastructure, which includes a discussion about driving range and chargers. The discussion also mentions various technical aspects of the driving experience, such as charging times and mileage. The most dominant topic discussed in Utah is the technical specifications of the vehicles, especially in Salt Lake City. The Utah topic model also indicates the presence of administrative factors, including taxes and the government.

Figure 8. Case study: Topic modeling and sentiment analysis results in Colorado vs. Utah.

Our case study suggests that users across both Colorado and Utah Reddit communities hold a positive opinion toward EV-related discussions in Figure 8a,b. The sentiment distribution indicates a similar curve. Despite Utah being a predominantly right-leaning state, public opinion toward EVs is extremely positive. This case study highlights the importance of location as a factor that affects public opinion, and we cannot rely solely on political affiliation/leaning to understand people’s opinions.
6. Discussion

Our study’s contributions include utilizing large-scale OSN datasets to analyze the demographic-based public perception of EVs. The methodologies employed, particularly in analyzing Twitter and Reddit posts, have enabled us to draw significant insights into the gender, age, and political orientation in the public opinion of EVs. This is evidenced by our detailed sentiment analysis and demographic categorization, which revealed the underrepresentation of certain groups, such as women and younger users, in online EV discussions.

Through the advanced analytical methods applied, we identified the prominence of certain topics, like charging infrastructure and tax benefits, across various demographic groups. These findings are a direct result of our methodical approach, which included sentiment analysis and topic modeling, thereby providing a data-driven understanding of public concerns and interests. The prominence of charging infrastructure issues, a critical factor identified in our analysis, highlights a widespread concern that transcends demographic boundaries and aligns with the technical challenges faced by EV adopters. This alignment between our methodology and the resulting insights underscores the effectiveness of our approach in capturing the pulse of public opinion.

Our findings also resonate with those of a previous large-scale customer survey [12], underscoring the validity of our results. The survey indicates that 75% of new EV buyers were male, closely aligning with our observation of a 74% share of voice among males in EV-related discussions on OSNs. Additionally, despite varying political affiliations, a generally positive sentiment toward EVs was prevalent, with a notably higher positive sentiment observed among Democrats compared with Republicans. This correlation between our gender and political analysis and external survey data not only validates our findings but also enriches the understanding of EV perception among OSN users, offering a comprehensive view that aligns with and complements previous large-scale surveys [12].

In reflecting on our methodology, we acknowledge that the choice of data sources and classification systems has shaped these findings. The reliance on social media data, while providing a broad and diverse set of opinions, also brings inherent biases related to the demographics of social media users. Furthermore, the accuracy of our demographic inferences, based on the sophisticated algorithms used, could have impacted the granularity of our results, a point we have critically examined in our approach.

Through this discussion, we aim to present a clear link between our methodological choices and the insights gleaned from our study. The findings demonstrate the effectiveness of our approach in capturing a diverse range of public opinions on EVs, highlighting the importance of demographic considerations in understanding and shaping the future of sustainable transportation.

7. Limitations and Future Work

This study, while providing insights into the demographic-based public perceptions of electric vehicles via online social networks, encounters several limitations. Primarily, the reliance on data from Twitter and Reddit introduces potential biases, as these platforms predominantly attract certain demographics, such as younger, more tech-savvy individuals, and may not fully capture the views of older or less digitally connected populations. Additionally, the classification systems used for demographic inference and political categorization are not without their shortcomings. Misclassifications, especially in nuanced demographic details or political leanings, could impact the findings. Moreover, the geographical focus of this study is limited, primarily drawing from users in the United States, which may not reflect global perspectives that vary significantly due to cultural, economic, and policy differences.

These limitations highlight the need for a more diverse and comprehensive approach in future research. Incorporating a broader array of data sources and expanding the geographical scope of the study could provide a more inclusive and accurate understanding of public opinion on electric vehicles. Continual refinement and validation of classification
algorithms against diverse datasets will also enhance the reliability of demographic and opinion analyses, contributing to more robust and inclusive sustainable transportation policies. There are several other recommendations for future work:

- OSN data can be collected over a long time range. Future research should consider the potential effects of temporal dynamics more carefully. For instance, different generations (e.g., Generation Y vs. millennials) may have different opinions on the same issue when they are young. Certain events can also have a significant influence on public perception. Associating temporal dynamics with real-world events can provide a deeper understanding of public perception.

- OSN data and traditional survey-based methods can be complementary. A combination of OSN discussions, surveys, or even news media reports can depict a more comprehensive picture of public perception of EVs in different demographic groups.

- Our study compares two states, i.e., Colorado and Utah; however, such comparisons can be extended to more states or even among different countries. Meanwhile, the political leaning comparison only takes left-leaning and right-leaning into consideration. Other political communities, such as centristm, can be examined in future work.

- Verified accounts exist widely on Twitter, and the political leaning of individuals, such as politicians, can provide valuable information to investigate how various political leanings perceive EVs differently. The results can be an essential complement to Reddit’s political comparisons in this study.

- In addition to age and gender, several other demographics play a role in the adoption of electric vehicles. These include the financial status of individuals, the household, geographical location, the number of cars already owned, and access to social media. Considering and analyzing all these factors can provide further insights into the perception of EVs.

- Future research could explore the evolving trends in topics of interest and sentiments toward EVs using long-term OSN data. Applying methods like the Structural Topic Model (STM) can reveal how public discussions and dominant topics about EVs have shifted over time. Additionally, analyzing sentiment trends across different time periods can provide insights into the changing nature of public opinion. This approach will enable a deeper understanding of how public interest and attitudes toward EVs evolve, reflecting the impact of market developments, technological advances, and environmental factors.

- An important direction for subsequent studies involves examining the impact of major commercial and political events on public perception of EVs. This could involve analyzing how specific events, such as the introduction of new EV models, significant policy announcements, or changes in government incentives, affect public discussions and sentiments. Utilizing causal inference methods, these studies could investigate whether there is a causal effect of certain policies on public perception. By correlating shifts in public opinion with these events, future research can provide valuable insights into the impact of specific policy interventions.

8. Conclusions

Through this research work, we have analyzed different demographic groups and how their opinions vary toward electric vehicles. We used the M3 system to classify the gender and age of Twitter users and conducted topic modeling analyses on all groups to understand their perceptions. It is important to note that the selection of the M3 system and topic modeling methodologies was driven by their established efficacy in large-scale social media data analysis. The M3 system is particularly valued for its accuracy in demographic inference, while topic modeling is effective in extracting salient themes from extensive text data. While a detailed discussion of the algorithms is beyond the scope of this work, we emphasize that these methodologies were chosen for their rigor and the transparent, replicable nature of their application, ensuring that our study contributes meaningfully to the broader field of EV perception research.
We also worked with data from Reddit to analyze how people with different political leanings respond to EV discussions. Our case study for Colorado and Utah demonstrates that only political leaning is not enough to understand the users’ behaviors. Adding other factors, such as financial and geographical, is also important.

The results of the analyses conducted in this work can be used to inform market strategies to better encourage people to adopt EVs. This information can also benefit policy-makers and help them understand the public’s opinion. For instance, given the under-representation of women and younger age groups in EV discussions, marketing strategies could be tailored to address the specific concerns and interests of these demographics, such as emphasizing environmental benefits or the long-term cost savings of EVs. Furthermore, our analysis of the prominence of charging infrastructure issues suggests a need for increased investment in and communication about the availability and convenience of EV charging stations, which could alleviate one of the major barriers to EV adoption. On the policy front, the positive sentiment we observed regarding tax benefits suggests that expanding financial incentives could be a powerful tool to encourage a wider range of demographics to consider EVs. Moreover, tailoring policies to accommodate the diverse opinions across political spectra, as revealed in our analysis, can make EV policies more inclusive and effective. Implementing these targeted interventions, informed by our data-driven insights, can lead to a more nuanced and successful approach to promoting EV adoption, ultimately supporting the transition to more sustainable transportation options.

In recognizing the limitations of this study, we acknowledge the potential biases and shortcomings inherent in our methodologies and data sources. The use of the M3 system for classifying gender and age on Twitter, while a useful tool, may not fully capture the entire spectrum of gender identities and could introduce inaccuracies. Similarly, our reliance on social media data from Twitter and Reddit may not encompass the full range of public opinions, as these platforms often attract specific demographics. Additionally, the simplification of complex political views into broad categories for analyzing Reddit data presents its own set of limitations. These factors underline the importance of interpreting our findings with an understanding of these constraints, and they highlight the need for future research to integrate more diverse data sources and refine classification methodologies. This approach will help achieve a more inclusive and accurate representation of public opinions on electric vehicles, contributing to a more nuanced understanding of the EV adoption landscape.

Author Contributions: Conceptualization, Q.L.; Methodology, T.R.; Validation, Q.L.; Formal analysis, T.P.; Investigation, T.P.; Data curation, T.R.; Writing—original draft, T.P.; Writing—review & editing, T.R. and Q.L.; Visualization, T.P.; Supervision, Q.L.; Project administration, Q.L.; Funding acquisition, Q.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported in part by the National Science Foundation (NSF) under Grant No. 1941524 for the NSF ASPIRE Engineering Research Center (ERC).

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References
29. Acar, A.; Muraki, Y. Twitter for crisis communication: Lessons learned from Japan’s tsunami disaster. *Int. J. Web Based Communities* 2011, 7, 392–402. [CrossRef]


43. Tusinski, D. Tusinski: From swing to safe, why Colorado is no longer a purple state. Rocky Mountain Collegian, 29 November 2022.


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.