Topic Modeling and Sentiment Analysis of Electric Vehicles of Twitter Data

H. P. Suresha[1*] and Krishna Kumar Tiwari[2]

[1]*REVA Academy for Corporate Excellence, REVA University Bengaluru, India.*
[2]*ML AI Community, Bengaluru, India.*

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Original Research Article

ABSTRACT

Twitter is a well-known social media tool for people to communicate their thoughts and feelings about products or services. In this project, I collect electric vehicles related user tweets from Twitter using Twitter API and analyze public perceptions and feelings regarding electric vehicles. After collecting the data, To begin with, as the first step, I built a pre-processed data model based on natural language processing (NLP) methods to select tweets. In the second step, I use topic modeling, word cloud, and EDA to examine several aspects of electric vehicles. By using Latent Dirichlet allocation, do Topic modeling to infer the various topics of electric vehicles. The topic modeling in this study was compared with LSA and LDA, and I found that LDA provides a better insight into topics, as well as better accuracy than LSA.In the third step, the "Valence Aware Dictionary (VADER)" and "sEntiment Reasoner (SONAR)" are used to analyze sentiment of electric vehicles, and its related tweets are either positive, negative, or neutral. In this project, I collected 45000 tweets from Twitter API, related hashtags, user location, and different topics of electric vehicles. Tesla is the top hashtag Twitter users tweeted while sharing tweets related to electric vehicles. Ekero Sweden is the most common location of users related to electric vehicles tweets. Tesla is the most common word in the tweets related to electric vehicles. Elon-musk is the common bi-gram found in the tweets related to electric vehicles. 47.1% of tweets are positive, 42.4% are neutral, and 10.5% are negative as per VADER Finally, I deploy this project work as a fully functional web app.

_____

*Corresponding author: E-mail: sureshhp.ba05@reva.edu.in;

1. INTRODUCTION

There are a variety of methods for researching and quantifying public opinion on electric vehicles. In order to accomplish this, Twitter is being used as a source of data in a social network study [1].

Social media platforms, such as Twitter, allow 280-character messages to be sent and received. Twitter's popularity has grown over time, and it is currently one of the most popular social media platforms [2].

Overall, social media is becoming increasingly popular, and it is now one of the primary means of communication for both individuals and businesses. As a result, tracking the flow of data on Twitter may assist you in keeping track of current events and comprehend how others are feeling [3].

Sentiment analysis and topic modelling are two analytical techniques that may be used to automatically handle Twitter data. They can be used for a variety of purposes, including event monitoring, product monitoring, and opinion mining. Companies do require timely and reliable information to react to market developments [4].

Topic modeling and sentiment analysis are two statistical techniques that can be effective in this situation. Topic modeling is any technique for uncovering a corpus' hidden semantic structure and gaining insight into the various themes contained in the texts [5].

In this project, I explored the topic modeling performance which uses the tweet's content to determine topics. LDA was used to evaluate the topics considering topic coherence in terms of their coherence [4].

The tweets data are positive, negative, and neutral sentiments on electric vehicles are determined using sentiment analysis. A text mining technique such as these could unlock great insights into the types of discussion happening on Twitter for electric vehicles [3] [5].

2. LITERATURE REVIEW

NLP is one of the most critical areas in social media since it deals with the interrelationships between large amounts of unstructured text. In NLP, the main goal is to understand how computer systems examine and extract information from human languages [6].

With the rise of social networks, several researchers have studied short text messages by applying sentiment analysis. Several studies have been conducted using Twitter samples. Twitter's API offers a convenient means to retrieve data [5].

Identify the topics and sentiment towards Electric vehicles by extracting Twitter user comments, by text pre-processing and NLP techniques [7].

The topic models used for inferring latent topics have been the focus of several studies recently. A topic model of an electric vehicle is usually modeled with LDA [4].

There are multiple opinions expressed about the data in a tweet, because different users express them in different ways.

"Pre-processing of tweet include following points -
Remove all URLs (e.g. www.xyz.com), hashtags (e.g. #topic), targets (@username)
Correct the spellings; sequence of repeated characters
Replace all the emoticons with their sentiment.
Remove all punctuations, symbols, numbers.
Remove Stop Words

Expand Acronyms (we can use an acronym dictionary)" [2].

Achieve Sentiment Classification by classifying a function as positive as well as negative or neutral. A corpus-based approach assigns each word to the emotional affinity and then identifies from the vast corpus to the probabilistic score. A system based on machine learning uses the technique of classifying text [2].

As the dataset contains various subjects of discussion, it is difficult to isolate them. It is

impossible to read all of the documents individually in a dataset this large given that each tweet represents one document. There are no labelled topics in the tweets. Topic modeling, an unsupervised machine learning technique, is employed to identify which topics are present and in what quantities. These topics are probabilistic mixtures of words that represent word co-occurrence trends in the dataset. Human-readable topics that are distinct from one another should be created by a good topic model [2].

The feature extraction method is used to extract characteristics from a processed dataset. Utilize later this feature to compute the polarity of a sentence, which aids the use of models such as unigrams and bigrams to determine individual opinions.

"Exciting features from processed tweet are 1.

1. Words and Their Frequencies: Unigrams, bigrams, and n-gram models with their frequency count
2. Parts of Speech Tags Parts of speech like adjectives, adverbs, and some groups of verbs and nouns are good indicators of subjectivity and sentiment.
3. Opinion Words and Phrases Apart from specific words
4. Position of Terms the position of a term within a text can affect the overall sentiment of the text.
5. Negation is an essential but complex feature to interpret. The presence of a negation usually changes the polarity of the opinion.
6. Syntax Syntactic patterns like collocations are used as features to learn subjectivity patterns."[2]

In this project, LDA is an algorithm used to train a topic model. A topic in LDA is simply the probability distribution of every word in the corpus. Word co-occurrences within documents are analyzed using LDA.

Latent Dirichlet allocation "(LDA): is a generative model, meaning that the model can be used to generate a corpus. Fig. 1 shows a plate diagram of LDA's productive process. Suppose our corpus has M documents, each with Nm (m from 1 to M) words with a total of distinct W words in the corpus overall, and suppose there will be k topics. α and η are hyperparameters for Dirichlet distributions that produce the k-dimensional document/topic (θ) vectors and W-dimensional topic/word (β) vectors, respectively. θ and β will be parameters for categorical distributions, will select topics and words (sampled). The generation process is as follows: 1. For each topic, j, βj is sampled. 2. For each document, m, θm is sampled. 3. For each word position n in document m (so this process is repeated Nm times), a topic z from the categorical distribution parameterized by θm. 4. Finally, a parameterized word from the categorical distribution by βz. Training an LDA model aims to determine θ and β that maximize generating the actual corpus. LDA takes α, η, and k as parameters and randomizes all other values (other than w). Then, each iteration slowly improves these values, using the guiding principles that words that occur in the same document are likely to be on the same topic, and documents that contain the exact words possible have some of the same topics. After many iterations, which obtain a fully trained LDA model, the main objects of interest being the document/topic and topic/word matrices [7]
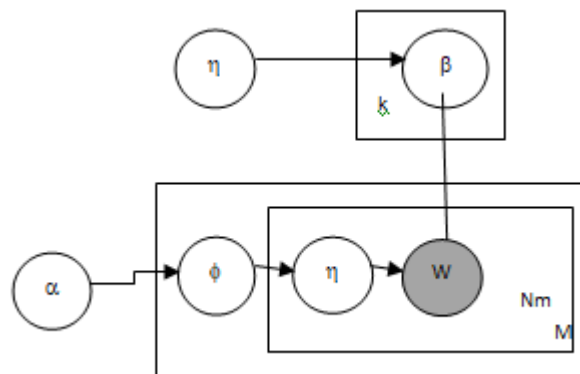

Fig. 1. The generative process of LDA

In Fig. 1 variables are represented by circles along with papers, words, and topics by rectangles (plates). In the corpus, only a shaded circle is visible; the rest are hidden in the model.

2.1 Natural Language Toolkit (NLTK)

NLTK is a Python package that includes several tools for data classification and programming.

It will be useful to linguists, engineers, students, educators, academics, and developers involved in natural language processing and text analytics. The NLTK library allows users to access over lexical resources with ease and 50 corpora. It consists of a collection of text processing packages for classification, tokenization, stemming, tagging, parsing, and reasoning" [8].

2.2 Valence Aware Dictionary and Entiment Reasoner (VADER)

The VADER sentiment analysis tool has a lexicon and rules tailored for the sentiment broadcast on social media. The program is open-source and free.

Word order and degree modifiers are also taken into account by VADER [8].

Development of NLP dashboard, web application and Cloud deployment API will serve as master source for the enterprises to identify the topic models and sentiments of electric vehicles and to take business decisions" [9-12].

3. OBJECTIVES OF THE STUDY

The main objective is to understand how electric vehicles and related topics and sentiments are trending in social media.

The first step is to collect tweets related to electric vehicles with their related hashtags with the help of Twitter API.

In this project, I like to perform EDA on the Twitter data of electric vehicles. From EDA, I want to collect the top 25 hashtags, maximum of 25 words, top 25 locations of users. This information will be helpful to know which are the hashtags, words, and places of users are trending in electric vehicles related tweets. EDA will bring a lot of value addition for the

Automotive manufacturers and suppliers to focus on these critical topics.

In this project, the next objective is to develop topic models using the LDA algorithm, which will be helpful to identify key topics of electric vehicles. It will be beneficial for the user to visualize the topics identified by LDA using suitable visual plots.

Understanding user sentiments towards electric vehicles will be beneficial to focus on the topics for the car manufacturers and suppliers. This project aims to implement VADER and SONAR sentiment analysis models to perform sentiment analysis of user tweets.

In this project, the final step is to develop a web application deployed in the cloud, which will be beneficial for the end-users and the organizations. [10] [13] [9]

The main objective of this project includes:

- Exploratory data analysis of tweets, hashtags, user location, and keywords.
- The development of the topic model to identify the key topics.
- The visualization of different topics from the electric vehicle tweets.
- Result of VADER sentiment analysis to identify positive, negative, and neutral tweets.
- Development of web application and deployment for the electric vehicle twitter data.

4. BUSINESS UNDERSTANDING

Compared to the 2020 forecast of 1.4 million units, the U.S. market for electric vehicles (EVs) could reach 6.9 million units by 2025. There is a short-term impact in the area of Electric vehicles due to pandemics; this will not stop any hindrance to the electric vehicle market and its development during the next 5 to 10 years. [1]

Automotive car manufacturers and suppliers expect to know user preferences to adopt new technology based on the end-user's requirements. Electric vehicles can help make these policies more efficient and effective. A full literature review would be helpful to summarise the findings and promote a more comprehensive knowledge of the topic because there have been a number of empirical research published on

consumer preferences for EV over the past decades [14].

There are many countries that have begun to implement policies to help grow electric vehicle (EV) production and adoption.

In the future development and success of emerging technologies, a significant amount of risk exists. As companies decide whether or not to accept new technologies, consumers' voices can help them lessen the uncertainty and reduce the risk. Our findings indicate that risk and benefit perceptions are important predictors of developing technology adoption [15].

In 2020, the pandemic had a significant impact on all types of automobile manufacturers. In the first half of 2020, new car registrations fell by about a third compared to the previous year. A 16 percent drop over last year was somewhat offset by more vital activity in the second half. The share of electric cars in global sales has increased by 70% to a record 4.6 percent in 2020, despite the decline of conventional and overall new car registrations [1].

The number of registered electric vehicles in 2020 is estimated to be around 3 million. For the first time, Europe took the lead with 1.4 million new registrations [1].

5. DATA ACQUISITION AND UNDERSTANDING

In this project, Acquired Tweets with the help of Twitter API about electric vehicles. I referred several journals and research papers about electric vehicles, Twitter topic modeling, Twitter sentiment analysis, Web application using Streamlit, and deployment in Heroku.

The following are the relevant dataset details:

The dataset consists of 45000 tweets, and its columns are as follows.

● Tweet Text: It refers to the text of the tweet
● Hashtag: Hashtag of the tweet
● Tweet datetime: the date and time of the tweet
● Twitter user location: It's the user's location

Table 1. Shows dataset sample from the acquired Tweets from Twitter

The Hashtags with user comments found in Twitter Tweets will be beneficial. A hashtag is a symbol; # is a metadata tag used on social networks like Twitter. The hashtag to tag the respective topic of discussion or opinion will help other users to find the related comments quickly with the hashtags.

Table 2 shows some of the hashtags related to Electric Vehicles from which the collected user comments are from Twitter.

Following are the activities during Twitter data preparation. After data collection/extraction, Data Processing is a crucial step to make the data ready for any modeling process:

Lower uppercase letters: The first step in pre-processing is to go through all of the data and convert every uppercase letter to its lowercase equivalent.

Remove URLs and user references: Twitter allows users to include hashtags, user references, and URLs in their messages. In most cases, user references and URLs are not relevant for analyzing the content of a text.

Remove digits: Numbers with Tweets aren't necessary for text data analysis. The same word with and without numbers creates a distinct meaning. example: iPhone and iPhone7.

Remove stop words: In natural language processing, the removal of stop words from the sample is standard. Removing stop words allows reducing the number of features extracted from the samples.

Remove repeated letters: This pre-processing step relates to the fact that when Twitter users want to highlight a topic, they frequently repeat particular characters numerous times.

Tokenize: Segmentation already happens in English, and tokenization is implicit. Because a space separates each word, it can make the token by dividing the text on each space.

Detect POS tags: For data analysis, a part of speech could be helpful in two ways. For instance, one can use it to clarify a word's meaning. The second application of POS tags is classifying terms and processing them differently depending on which kind they belong.

Table 1. Electric vehicles tweets data sample

Tweet Text	Hashtag	Tweet Datetime	Twitter User Location
Almost ready - first phase of the largest electric vehicle factory in the world: Aggarwalâ€¦ https://t.co/7bFmMZ5Cea	Electricvehicle	6/27/2021 8:55	Mumbai, India
RT @gripinvest: What's the first-ever investment you've made? Let us know in the comments below!ðŸ',¸ðŸ¤"ðŸ'‡ Visit https://t.co/e3Bfrz6VS6 to viewâ€¦	Electricvehicle	6/27/2021 8:31	New Delhi, India
RT @BuopsoNews: MG Motor to drive in their 2nd electric model in 2 years @MGMotorIn #ElectricVehicle #business #tech #innovation #news #Suâ€¦	Electricvehicle	6/27/2021 8:22	Palo Alto, CA
MG Motor to drive in their 2nd electric model in 2 years @MGMotorIn #ElectricVehicle #business #tech #innovationâ€¦ https://t.co/PcngbJmdbb	Electricvehicle	6/27/2021 8:15	Gujranwala Town, Delhi, 110009

Table 2. Major Hashtags from Electric vehicles tweets

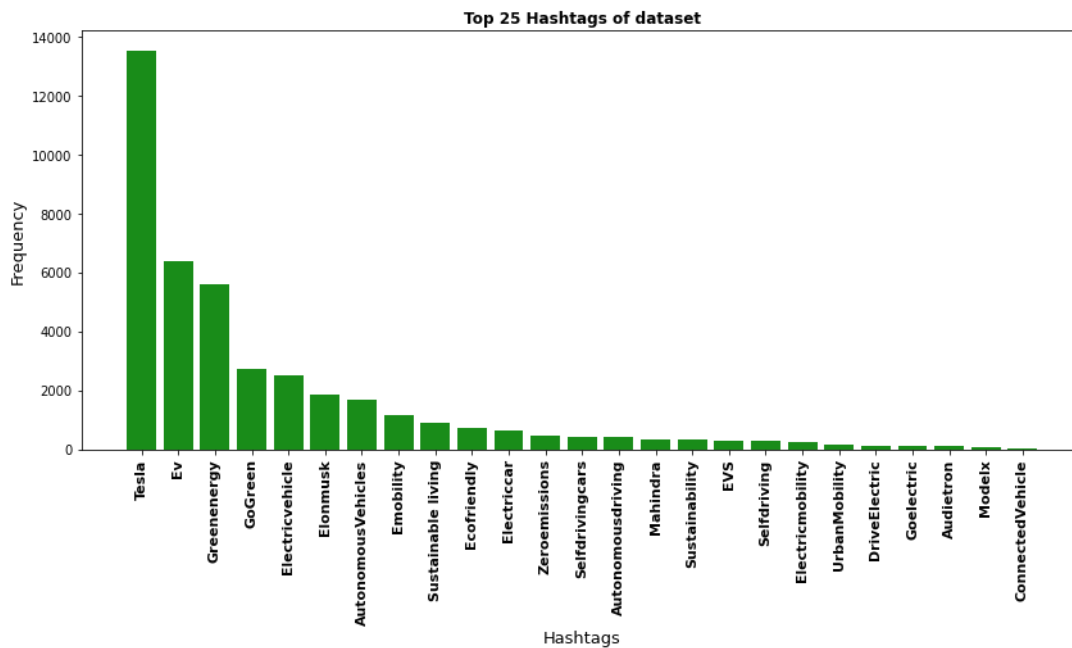| #Alternative fuel	#Audietron	#Autonomousdriving
#AutonomousVehicles	#Carcharger	#ConnectedVehicle
#DriveElectric	#Ecofriendly	#Electriccar
#Electricfuture	#Electricmobility	#Electricvehicle
#Electricevehiclecharging	#Elonmusk	#Emobility
#Ev	#EvBAttery	#Evconversion
#EVS	#EVSales	#Futurecars
#Goelectric	#GoGreen	#Greenenergy
#Mahindra	#Modelx	#Selfdriving
#Selfdrivingcars	#Sustainability	#Sustainable living
#Tesla	#Teslacars	#Teslacybertruck
#Teslalife	#TeslaMotors	#Teslaowner
#Teslaroadster	#UrbanMobility	#Zeroemissions

Lemmatize: When processing samples, "word" and "words" would be considered as two different features. Hence, to improve the features reduction process, the unigrams can be lemmatized. This pre-processing step mainly allows removing plurals and conjugations" [7].

Fig. 2 shows the top 25 Hashtags from the Twitter data collected related to electric vehicles. Tesla is the top Hashtags most commonly used in the Tweets.

Fig. 3 shows the top 25 user locations from the Twitter data collected related to electric vehicles. Sweden, India and USA are the top 3 user locations most users Tweeted in Twitter.

Fig. 4 shows the word cloud from the Twitter data collected related to electric vehicles. Tesla, future, electric vehicle, Elon musk are the most frequent words on Twitter.
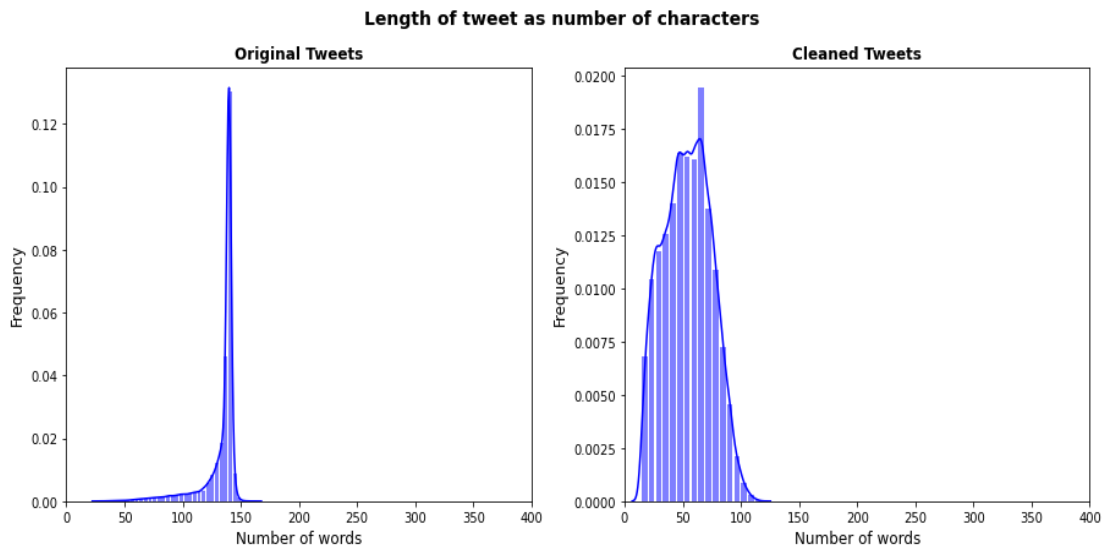
Fig. 5 shows the consequences of the pre-processing phase significantly decreased the length of cleaned tweets. After pre-processing, most tweets include less than ten tokens, as illustrated, compared to roughly 20 words in original tweets. The impacts are more evident in the second set of graphs, where the tweet length decreased from around 150 to around 100 characters following the cleaning stage. This phase is crucial since it reduces dimensionality.

Fig. 2. Top 25 Hashtags from Electric vehicles tweets

Fig. 3. Top 25 locations of Electric vehicles tweets

Fig. 4. Word cloud of Electric vehicles tweets

Length of tweet as number of characters

Fig. 5. Pre-processed and processed cleaned Electric vehicles tweets

Fig. 6 shows the top 25 words from the Tweets after pre-processing. Tesla, project and electric are the top most common words.

6. DATA MODELING

I started topic modeling with the processed and clean tweets data from electric vehicles. The purpose of topic modeling is to find hidden subjects in large amounts of text. The Latent Dirichlet Allocation (LDA) technique is one of the standard topic modeling algorithms of Python's Gensim package. A LDA topic model makes use of the dictionary(id2word) and the corpus as inputs [16].

Latent Dirichlet Allocation (LDA) via Gensim will be used in conjunction with Mallet's implementation (via Gensim). Mallet has a well-functioning LDA implementation. Mallet helps LDA to be faster and provides greater topic separation.

Bigrams are two words that commonly appear in the same sentence in a document. Trigrams are groups of three words that occur frequently. Elon-musk, electric-vehicle are some top bigrams from our data.

The Phrases model from Gensim can create and implement bigrams, trigrams, and n-grams. Among Phrases' parameters, minimum and threshold are two of the most important. In general, the larger these parameters are, the harder it is to combine words into bigrams.

Fig. 7 shows top 25 Bigrams identified from electric vehicles tweets from topic modeling approach, elon-musk, electric-vehicle are most common Bigrams.

To illustrate the top 20 frequently occurring bigrams, I used the Python program NetworkX. Fig. 8 shows most frequent bigrams related to electric vehicles.

Let's visualize the subjects for interpretability now that I have a trained model. To accomplish so, we'll utilize pyLDAvis, a popular visualization software aimed to aid in interactively better understanding and to analyze individual topics and better experience the links between them.

1. By manually selecting each topic with alternative λ parameters, we can see the most common and relevant phrases for that topic.

2. Investigating the Intertopic Distance Plot will help you understand how topics are related to one another, as well as possible higher-level structure.

Fig. 9 shows interactive Topic modelling visualization plot from LDA from Electric vehicles tweets.

The degree of semantic similarity among terms in a topic is measured by topic coherence. As a result of these metrics, semantically interpretable issues can be distinguished from statistically inferable ones. Fig. 10 shows that

as the number of topics grows, the coherence score rises. Now, the number of subjects you choose is still determined by your needs, as topics with a coherence score of approximately 25 have strong coherence but may contain repeated keywords.

In VADER sentimental analysis, lexical features are mapped to emotion intensities known as sentiment scores. The sentiment score can be calculated by adding together the word intensities within a text.
With VADER, you can input a string and get a dictionary of scores in four categories:

- negative
- neutral
- positive [17]

Fig. 11 shows distribution of positive, negative and neutral tweets identified from VADER sentiment analysis model.

VADER gives a compound score of whether a sentence leans in the direction of positive, neutral, or negative. Compound score work as follows:

- Using the words in the lexicon, sum the scores, adjust according to the rules (all files), and scale to -1 to +1.
- VADER calculates a compound score using ratios from the portions of text in each category.
- A positive result is greater than 0.05
- In the neutral range of -0.05 and 0.05

- Negative values are less than or equal to -0.05

Fig. 12 shows density plot of overall compound score, we have more positive tweets related to electric vehicles compared to negative and neutral.

Fig. 13 shows Word cloud of positive sentiment, it is identified from VADER

Fig. 14 shows Word cloud of negative sentiment, it is identified from VADER

I used the sonar python library in python to identify Hate and offensive tweets related to electric vehicles as follows:
Fig. 15 shows bar chart of hateful and offensive tweets. It is identified from SONAR library.

Most of these hateful and offensive tweets are about tesla decision about non acceptance of bit coin.

Specifically, I used my trained LDA model to determine the topic composition of each sentence. In a sentence, if one topic is dominated by more than 70%, I considered that sentence to belong to that topic. After that, I calculated the sentiment of the sentence using Vader as positive, negative, or neutral, and finally counted the total percentage of each topic for the final split between positive, negative, or neutral.

7. DEPLOYMENT

I constructed a basic web application with a form field that allows users to enter information, examine charts and results using streamlit, stored and maintained essential files in GitHub, and launched it on Heroku for this project. Streamlit is a data science-oriented application framework. In this project, I used the Streamlit framework to develop a functional web app. Deployment is done using the Heroku cloud platform [10] Heroku From their own words: Heroku makes it easy to build, deploy, monitor, and scale apps, eliminating time consuming infrastructure headaches [16].

In addition to hosting Git repositories, GitHub offers many of its own features. In contrast to Git, which is a command-line tool, GitHub is a graphical user interface on the Web. In addition, it offers access control and several collaboration tools, such as wikis and task management tools, for every project. [13] Fig. 16 shows how the web application, GitHub and Heroku cloud interacting during deployment and end users can able to use this application to analyze tweets. Web link of this application is as follows: https://nlp-sureshhp.herokuapp.com/ Fig. 16 Web app deployment.

8. ANALYSIS AND RESULTS

In this project, I collected 45000 tweets from Twitter API, related hashtags, user location, and different topics of electric vehicles.

Tesla is the top hashtag Twitter users tweeted while sharing tweets related to electric vehicles. Ekero Sweden is the most common location of users related to electric vehicles tweets.

Tesla is the most common word in the tweets related to electric vehicles. Elon-musk is the common bi-gram found in the tweets related to electric vehicles. 47.1% of tweets are positive, 42.4% are neutral, and 10.5% are negative as per VADER.

SONAR identified critical tweets which were hateful and offensive. These tweets happened during tesla declared nonacceptance of bitcoin for the purchase of electric vehicles.

Finally, I built topic models, and Vader and sonar sentiment analysis models created a Streamlit web app and deployed it in Heroku through the GitHub repository.

Functionalities added in the web application are as follows:
Sample dataset

25 top hashtags plot
25 Most frequent words
Vader sentiment plot
Enter text Box – enter new tweets or text
Tokenize
Pos Tagging

Translate (60 languages)
Sentiment analysis
Summarize
Display n-grams, Specific word count and Word frequency visualization

Limitations of LDA and VADER: Latent Dirichlet allocation models cannot identify Corelation, Static cannot topics identify over change in time, VADER is not accurate for complex data, does not recognize context, required additional visualization output to understand the results.
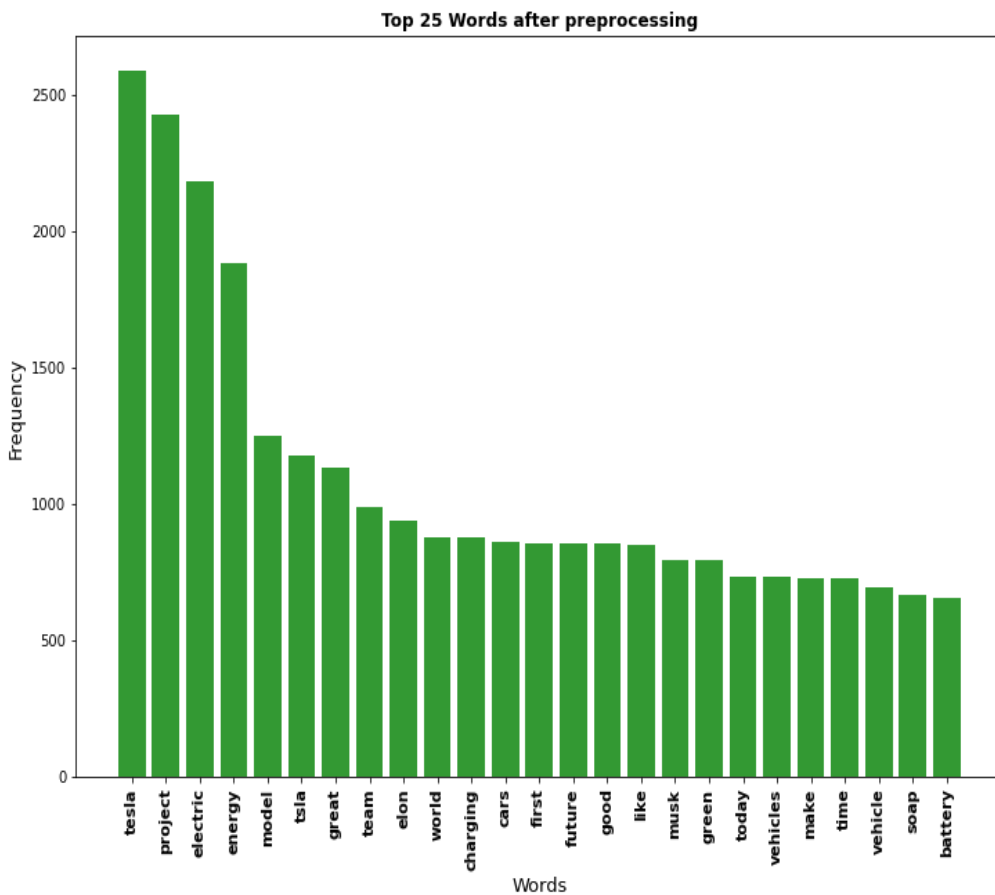

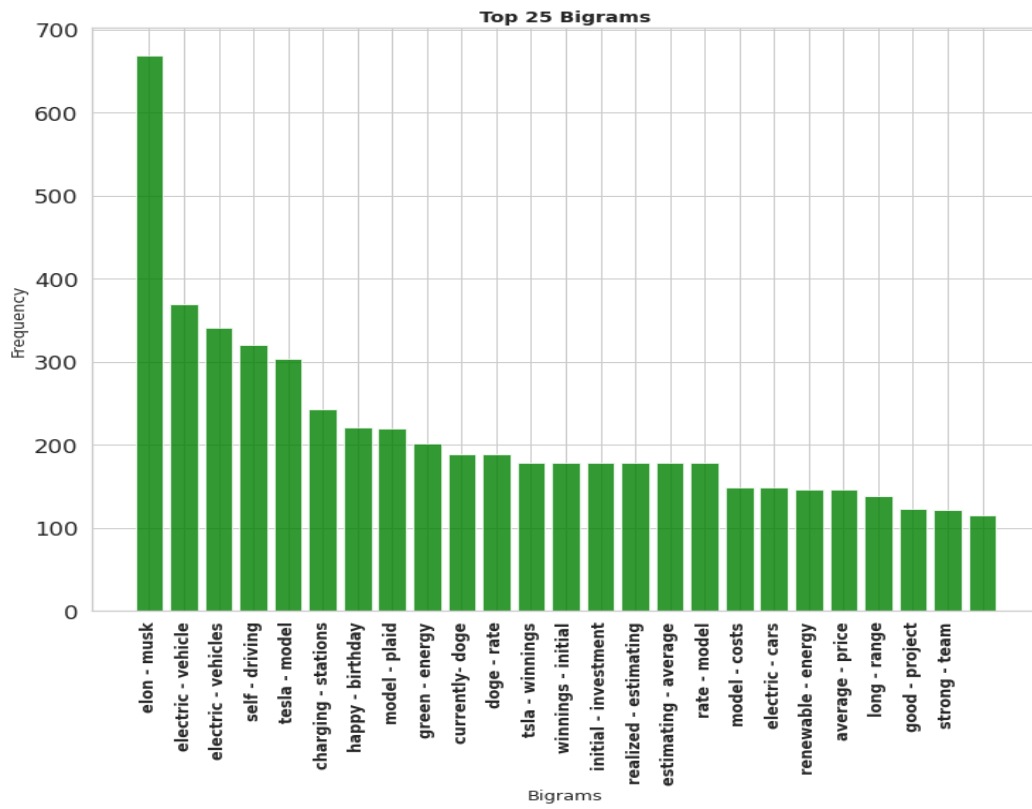**Fig. 6. Top 25 words of Electric vehicles tweets after pre processing**

Fig. 7. Top 25 Bi Grams words from Electric vehicles tweets
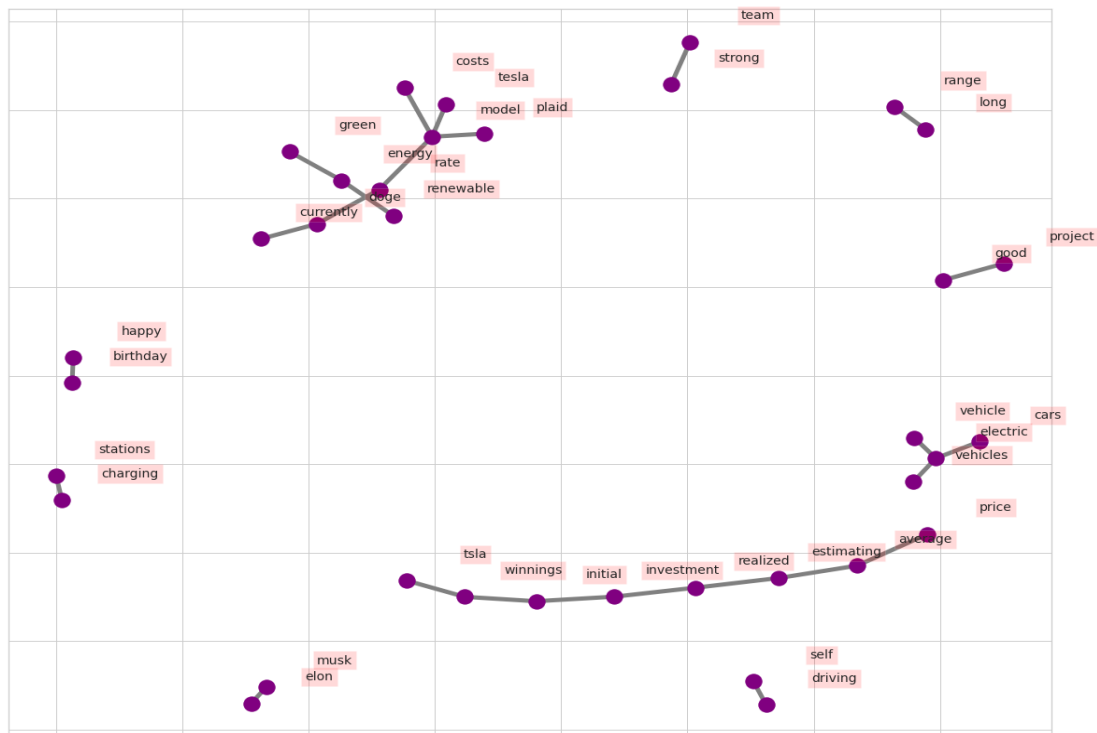

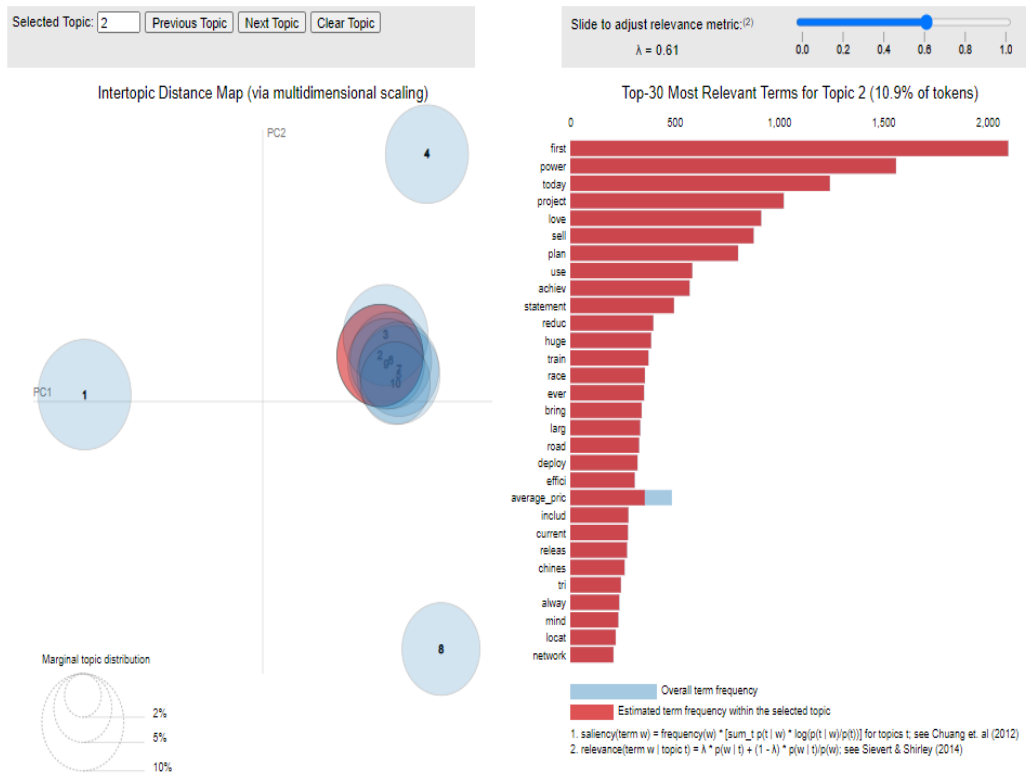**Fig. 8. Visualization of bigrams from Electric vehicles tweets**

Fig. 9. Visualization Topic modelling from LDA from Electric vehicles tweets
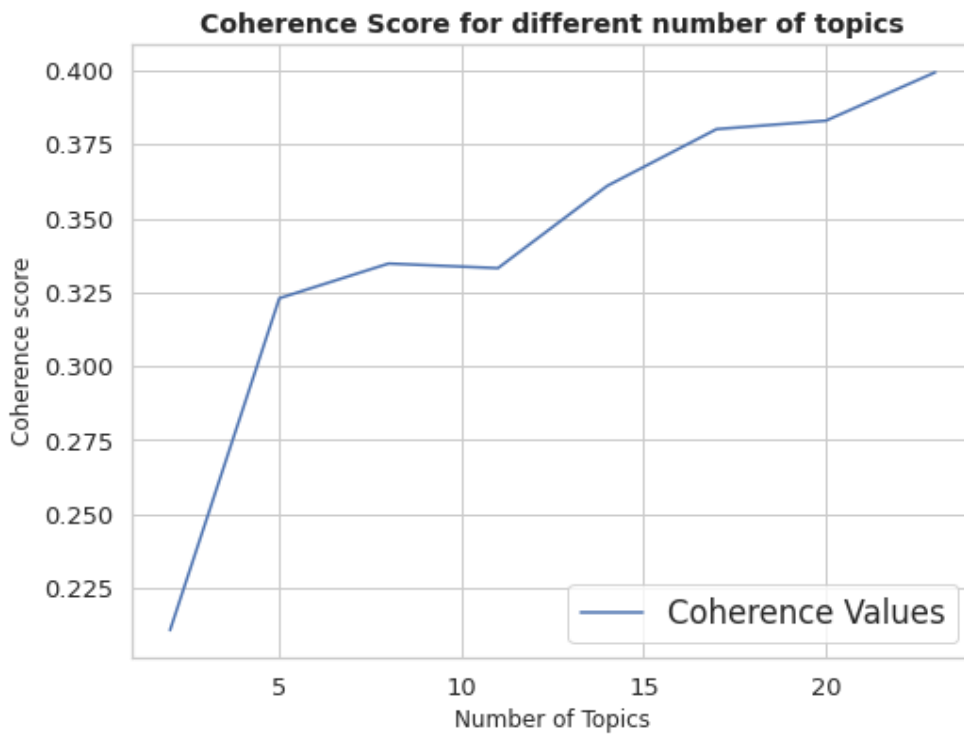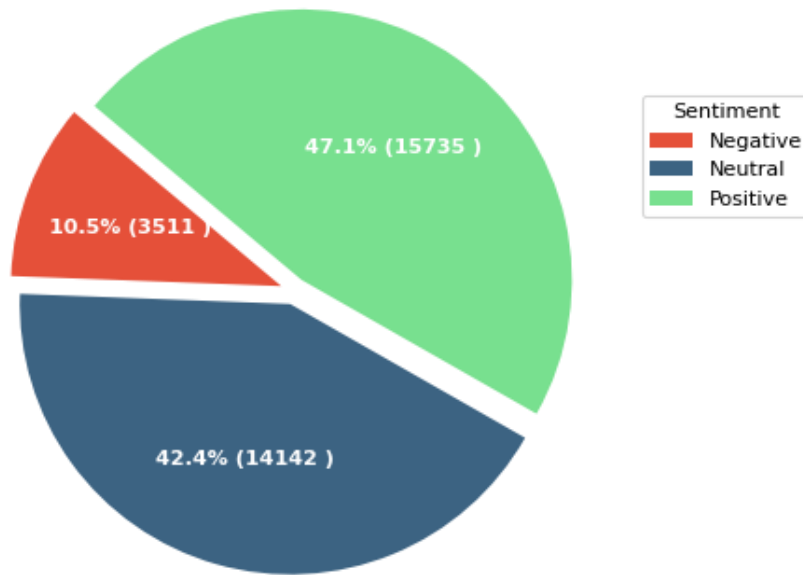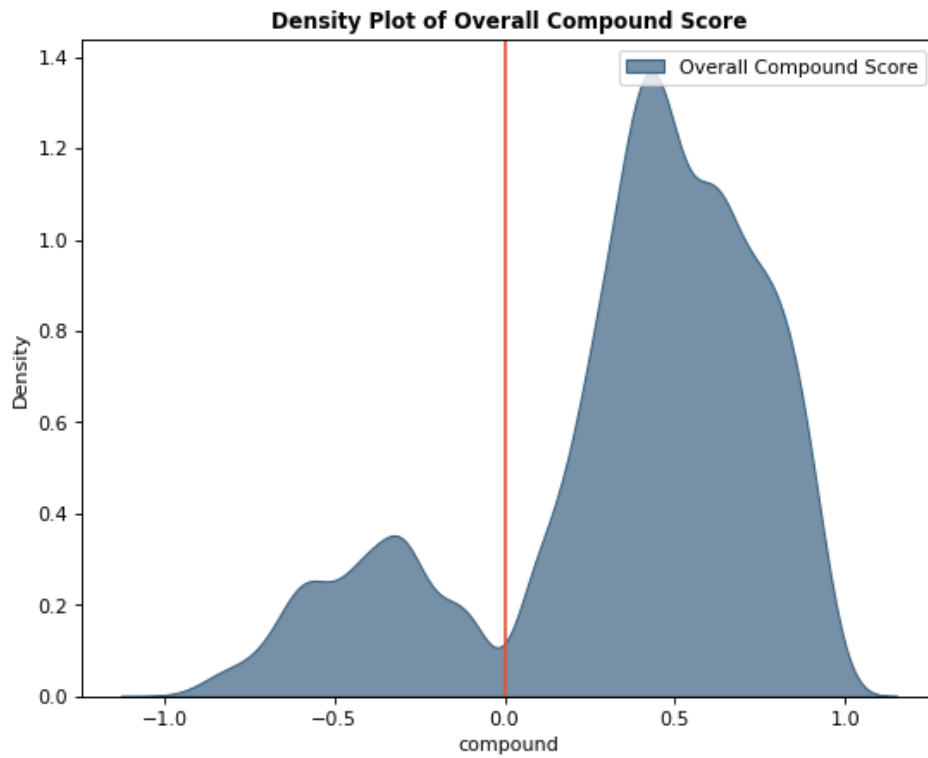

Fig. 10. Visualization coherence score based on topics from LDA

Number of Tweets by Sentiment

Fig. 11. Number of tweets by Sentiment

Fig. 12. Density Plot of overall compound score

WordCloud of Positive Tweets

Fig. 13. Word cloud of positive sentiment

WordCloud of Negative Tweets

Fig. 14. Word cloud of negative sentiment

Fig. 15. Bar chart for Hateful and offensive tweets

Fig. 16. Web app deployment

9. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

Twitter's discussion about electric cars informed us about significant events regarding electric cars and mirrored mainstream media stories. A global perspective was also evident in the tweets discussion. According to social media discussion concerning electric vehicles, there is a growing sentiment of trust and positive sentiment indicating that electric vehicles will be more widely accepted in the near future.

In this project, I identified critical topics from topic modeling for electric vehicle gets using Twitter user comments. We have used various approaches of data science to extract and study topics and sentiments of users.

My crucial contribution in this project is to implement topic modeling to identify key topics, visualization using LDA, bi-grams, word cloud, dominant topics, VADER sentiment, and SONAR which will help in the future sentiment towards electric vehicles in India.

For the future scope, I continue to work on dynamic topic modeling concerning time and introduce multiple NLP classifiers to understand the accuracy of sentiment predicted along with VADER and SONAR. Apply additional functionalities in the web application.

In conclusion, this research gives us a better grasp of electric vehicles topics and sentiment in general, as well as specific subareas and scopes. It's worth noting that topic definitions and concepts might shift throughout time. The dynamic topic modeling of electric vehicles is not included in our analysis. Currently from the project's web application and methodological findings, all identified topics, sentiment, and geographic region are included.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Global EV Outlook 2020," Glob. EV Outlook; 2020.
    DOI: 10.1787/d394399e-en.
2. VA, Sonawane SS. Sentiment Analysis of Twitter Data: A Survey of Techniques," Int. J. Comput. Appl. 2016;139(11):5–15.
    DOI: 10.5120/ijca2016908625.
3. Alsaeedi A, Khan MZ. A study on sentiment analysis techniques of Twitter data," Int. J. Adv. Comput. Sci. Appl. 2019;10(2):361–374.
    DOI: 10.14569/ijacsa.2019.0100248.
4. Alash HM, Al-Sultany GA. Improve topic modeling algorithms based on Twitter hashtags," J. Phys. Conf. Ser. 1660;1:2020.
    DOI: 10.1088/1742-6596/1660/1/012100.
5. Mittal A, Patidar S. Sentiment analysis on twitter data: A survey," ACM Int. Conf. Proceeding Ser. 2019;91–95.
    DOI: 10.1145/3348445.3348466.
6. Salloum SA, Al-Emran M, Monem AA, Shaalan K. A survey of text mining in social media: Facebook and Twitter perspectives," Adv. Sci. Technol. Eng. Syst. 2017;2(1):127–133.
    DOI: 10.25046/aj020115.
7. Ficamos P, Liu Y. A Topic based Approach for Sentiment Analysis on Twitter Data," Int. J. Adv. Comput. Sci. Appl. 2016;7(12):201–205.
    DOI: 10.14569/ijacsa.2016.071226.
8. Elbagir S, Yang J. "Twitter sentiment analysis using natural language toolkit and Vader sentiment," Lect. Notes Eng. Comput. Sci. 2019;2239:12–16.
9. Heroku, "Heroku," 2021, [Online].

Available:https://devcenter.heroku.com/articles/how-heroku-works.
10. Streamlit, "Streamlit; 2021, [Online].
    Available:https://docs.streamlit.io/en/stable
11. Shukri SE, Yaghi RI, Aljarah I, Alsawalqah H. Twitter sentiment analysis: A case study in the automotive industry," 2015 IEEE Jordan Conf. Appl. Electr. Eng. Comput. Technol. AEECT; 2015.
    DOI: 10.1109/AEECT.2015.7360594.
12. Jesse E.Agbe, "Streamlit app; 2021.
    Available:https://blog.jcharistech.com/2019/10/22/building-a-natural-language-processing-app-with-streamlitspacy-and-python/.
13. Github, Github; 2021, [Online].
    Available:https://docs.github.com/en/github
14. Liao F, Molin E, Van Wee B. "Consumer preferences for electric vehicles: a literature review," Transp. Rev. 2017;37(3):252–275.
    DOI: 10.1080/01441647.2016.1230794.
15. Kohl C, Knigge M, Baader G, Böhm M, Krcmar H. Anticipating acceptance of emerging technologies using twitter: the case of self-driving cars," J. Bus. Econ. 2018'88(5):617–642.
    DOI: 10.1007/s11573-018-0897-5.
16. Prabhakaran S. Topic modeling; 2018, [Online].
    Available:https://www.machinelearningplus com/nlp/topic-modeling-gensim-python/.
17. Beri A. Sentiment Analysis; 2020, [Online].
    Available:https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664.

_____