CLA4: Course Group Project Demo

Team Members

Yaswanthraj SP (RA2011027010003)

Sai Pradyumnan Prem (RA2011027010041)

Topic: Compare different types of vehicle's Performance, Range, Ride, Handling, Braking, Steering, Off-road capability, NVH, Occupant Comfort (HVAC, Seating), Visibility & Package, Safety, Software Features.

I. Introduction

A. Background and Motivation: This project aims to compare the performance, range, ride, handling, braking, steering, off-road capability, NVH, occupant comfort, visibility & package, safety, and software features of six different types of vehicles based on sentiment analysis of reviews scraped from multiple sources. The purpose is to provide insights into the strengths and weaknesses of these vehicles in various aspects and help potential buyers make informed decisions.

B. Objectives:

i. Collect and preprocess review data from multiple sources.

ii. Identify keywords related to the 12 performance parameters and score them using VADER lexicon.

iii. Analyze and compare the performance of the six vehicles in each parameter.

iv. Evaluate the occupant comfort, safety, and software features of the six vehicles.

v. Visualize and discuss the findings.

C. Scope and Limitations: The project will focus on six types of electric vehicles: Tata Tigor EV, Hyundai Kona, Kia EV6, Mercedes Benz EQ6, MG ZS EV, and Tata Nexon. The sentiment analysis will be based on reviews scraped from online platforms such as CarDekho, ZigWheels, and CarWale. The project will be limited to the English language only.

D. Methodology: The project will use sentiment analysis and VADER lexicon in the NLTK library to process the review data. The collected data will be preprocessed to remove noise and extract relevant keywords. The keywords will be scored based on their sentiment using VADER lexicon. The scores will be analyzed and compared across the six vehicles in each of the 12 performance parameters.

II. Data Collection and Preprocessing

A. Source of Data: The project will collect review data from online platforms such as CarDekho, ZigWheels, and CarWale.

B. Preprocessing: The collected data will be preprocessed to remove noise and extract relevant keywords related to the 12 performance parameters.

III. Keyword Identification and Scoring

A. Keyword Identification: The relevant keywords will be identified using the preprocessed review data.

B. Scoring: The keywords will be scored based on their sentiment using VADER lexicon.

IV. Data Analysis

A. Performance Comparison: The scores of each parameter across the six vehicles will be analyzed and compared.

B. Occupant Comfort, Safety, and Software Features: The occupant comfort, safety, and software features of the six vehicles will be evaluated.

V. Results and Discussion

A. Performance Comparison: The performance of the six vehicles in each parameter will be compared and discussed.

B. Occupant Comfort, Safety, and Software Features: The evaluation of occupant comfort, safety, and software features will be discussed.

C. Findings: The findings will be discussed and compared with the literature review.

Purpose (Usage) of the Proposed Project:

The purpose of this proposed project is to compare and evaluate the performance, range, ride, handling, braking, steering, off-road capability, NVH, occupant comfort, visibility & package, safety, and software features of six different types of vehicles based on the sentiment analysis of reviews scraped from multiple sources. The project aims to provide insights into the strengths and weaknesses of these vehicles in various aspects and help potential buyers make informed decisions. Moreover, the project also seeks to contribute to the field of sentiment analysis by demonstrating the use of VADER lexicon in processing large amounts of textual data and extracting meaningful information. Overall, the proposed project can benefit both vehicle manufacturers and buyers by facilitating product development and selection.

Explanation about Our Work

In the project mentioned above, we have chosen sentiment analysis using the VADER lexicon in the NLTK library to evaluate the performance, range, ride, handling, braking, steering, off-road capability, NVH, occupant comfort, visibility & package, safety, and software features of six different types of electric vehicles. We chose this model because sentiment analysis is an effective way to process textual data and extract valuable insights. The VADER lexicon is specifically designed to handle sentiment analysis of social media text and has been shown to provide accurate results in several studies. Additionally, the NLTK library provides a robust set of tools for natural language processing and has been widely used in research and industry. Overall, we believe that using sentiment analysis with the VADER lexicon in the NLTK library is an appropriate choice for our project.

Performance Analysis:

In a sentiment analysis project, accuracy can be evaluated by comparing the predicted sentiment (positive, negative, or neutral) of each review with the actual sentiment expressed in the review.

There are several metrics that can be used to evaluate the accuracy including precision, recall, F1 score, and accuracy. Precision measures the proportion of predicted positive reviews that are actually positive, recall measures the proportion of actual positive reviews that are correctly identified as positive, F1 score is the harmonic mean of precision and recall, and accuracy measures the proportion of all reviews that are correctly classified.

The accuracy of a sentiment analysis model can vary depending on several factors, including the quality and quantity of training data, the choice of algorithm, and the complexity of the model. Therefore, it's important to carefully evaluate and fine-tune the model to ensure high accuracy.

Conclusion:

The findings of this project will be useful for potential buyers of electric vehicles who are looking for objective and data-driven evaluations of the performance, comfort, safety, and software features of different types of vehicles. The analysis of the 12 parameters will provide insights into the strengths and weaknesses of each vehicle and help buyers make informed decisions based on their specific needs and preferences. Additionally, the project will be useful for vehicle manufacturers who are interested in understanding customer perceptions and preferences and improving their products accordingly. Finally, the project will be useful for researchers and practitioners in the field of sentiment analysis and natural language processing who are interested in exploring the application of these techniques in the domain of electric vehicles.