Market segmentation of EV market in India

By:- Jivesh Sharad Bharambe

GitHub:- [Link](#)

1. Problem Statement:-

The global push for sustainability and reduced carbon emissions has accelerated the adoption of electric vehicles (EVs). However, the EV market faces challenges in understanding diverse customer preferences, geographic adoption rates, and barriers to entry. The objective of this study is to:

- Identify key market segments within the EV domain.
- Understand factors driving EV adoption and barriers hindering growth.
- Provide actionable insights for targeted marketing and product development.

By addressing these goals, businesses can better cater to distinct customer groups and improve EV market penetration.

2. Data Collection

- Data Sources:
  The dataset used for this analysis was collected from multiple sources, including:
- Surveys:
  - Customer feedback forms covering demographics, income, and EV preferences.
  - Responses from urban and semi-urban regions.
- Publicly Available Datasets:
  - Market reports on EV adoption trends.
  - Open datasets containing regional charging station data.
- Company Data:
  - Sales and customer interaction data from EV manufacturers.
- Web Scraping:
  - Online reviews and ratings of EV models.
- Key Variables in the Dataset:
  - Demographics: Age, gender, marital status, profession, and annual income.
  - Geographic Data: City, state, and region.
  - Behavioral Data: EV preferences, daily mileage, and vehicle usage.
  - Product Preferences: Desired features, charging preferences, and willingness to pay.
  - Environmental Awareness: Scores indicating concern for sustainability and eco-friendliness.

➢ Data Cleaning:
  - Removed irrelevant columns (e.g., Unnamed: 0).
  - Handled missing values by imputing or dropping incomplete records.
  - Standardized categorical variables, such as city names, to ensure consistency.

With this robust dataset, the analysis could uncover significant patterns and insights to address the problem statement effectively.

3. Explain how and which ML model (algorithm) helped you in the 2nd Project:-

The project employed Decision Tree Classifier and K-Nearest Neighbors (KNN) algorithms for analysis and segmentation tasks. These supervised machine learning models were chosen due to their effectiveness in classifying and predicting customer segments based on historical data. Here is a detailed explanation of their roles:

Decision Tree Classifier:-

➢ Objective:- To build an interpretable model that identifies key decision points influencing customer segmentation.

Key Steps:-

- Features such as age, annual income, and EV preference were used to split the dataset into distinct groups.
- Each node in the tree represented a decision criterion (e.g., income > $50,000).
- The classifier provided a visual representation of how different variables influenced customer behavior.

Advantages:-

- High interpretability, allowing easy communication of segmentation logic to stakeholders.
- Ability to handle both categorical and numerical data.

K-Nearest Neighbors (KNN):-

➢ Objective: To classify new customers into predefined segments based on similarity.

Key Steps:-

- Standardized the dataset to ensure all features had equal weight in distance calculations.
- For each data point, the model identified the k nearest neighbors and assigned it to the segment with the majority class.
- Tested different values of k to optimize model performance.

Advantages:-

- Non-parametric nature, making it flexible to different data distributions.
- Effective in handling complex, non-linear relationships between features.

Combined Benefits

By using these models together, the project achieved a balance of interpretability (Decision Tree) and accuracy in classification (KNN). These models identified clear customer profiles and improved the understanding of market dynamics.

4. Elaborate on the final conclusion & insights gained from the research/analysis work:-

Key Findings:

- Distinct Customer Segments:-
  - Young Professionals: Predominantly prefer compact, cost-efficient EVs with modern designs.
  - High-Income Buyers: Favor luxury EVs with premium features such as extended range and advanced technology.
  - Environmentally Conscious Families: Opt for affordable EVs emphasizing eco-friendliness and sustainability.

- Geographic Preferences:-
  - Cities like Pune and Bengaluru exhibited higher interest in EVs, likely due to better infrastructure and urban initiatives.
  - Rural and semi-urban areas showed slower adoption rates.

- Critical Features Influencing Decisions:-
  - Charging station availability significantly impacted purchase preferences.
  - Aesthetic features, such as wheel design, played a role in influencing premium customers.

Business Implications:

The insights can drive customized marketing strategies and inform product development tailored to each segment's unique needs.

5. How will you improve upon the Market Segmentation Project given additional time & some budget to purchase data?

Enhancements in Data Collection

To refine the analysis, additional data points should be collected:

- **Behavioral Data**:
    - Daily mileage patterns.
    - Frequency of long-distance travel.

- **Infrastructure Data**:
    - Accessibility and density of charging stations in key markets.

- **Psychographic Data**:
    - Environmental awareness scores.
    - Lifestyle indicators (e.g., fitness-consciousness, technology adoption).

- **Competitor Analysis**:
    - Pricing and feature comparison of EVs from competing brands.
    - Customer reviews and satisfaction levels with competitors.

Advanced ML Models to Explore

- **Random Forest Classifier**:
    - To improve decision-making accuracy by aggregating results from multiple decision trees.

- **Support Vector Machines (SVM)**:
    - To handle non-linear relationships and identify clear decision boundaries.

- **Clustering Techniques**:
    - **Hierarchical Clustering**: For understanding hierarchical relationships within segments.
    - **DBSCAN**: To identify non-spherical clusters and outliers.

Additional Tools:

- Integration of **Principal Component Analysis (PCA)** for dimensionality reduction to identify the most impactful features for segmentation.

These enhancements will enable more precise segmentation and actionable insights.

6. What is the estimated Market Size for your Market Domain (non-segmented) in Numbers?

Current Market Estimates:

- **Global EV Market Size**:

    o Valued at approximately **$250 billion in 2023**.

    o Expected to grow at a **CAGR of ~24%**, reaching ~$800 billion by 2030.

- **Indian EV Market Size**:

    o Estimated at **$3 billion in 2023**, with substantial growth driven by government subsidies and urban adoption.

    o Projected to reach ~$20 billion by 2030.

Refinement Potential:

Further market size estimation can be enhanced by:

- Incorporating localized data on EV sales and registrations.

- Analyzing policy-driven growth in emerging markets.

7. Name top 4 Variables/features which can be used to create the most optimal Market Segments for your Market Domain.

Optimal Features for Segmentation:

1. **Annual Income**:

    o A critical factor influencing budget constraints and preferences for premium or budget EVs.

2. **Age**:

    o Directly correlates with technology adoption and design preferences.

3. **Preference for Wheels in EV**:

    o Reflects aesthetic and functional priorities, especially among premium buyers.

4. **City/Geographic Location**:

    o Indicates regional interest and accessibility to EV infrastructure, significantly impacting adoption rates.

By focusing on these features, the segmentation process can yield actionable and practical customer insights, ensuring tailored strategies for market success.