ISSN 2776-8007

Predicting Future Electric Vehicle (EV) Sales: A Time Series Forecasting Approach Using Historical EV Sales Data

Bhavana Srinivasan[1,*]

[1]*Department of Animation and Virtual Reality, JAIN, Bangalore, India*

(Received: March 13, 2025; Revised: April 15, 2025; Accepted: August 2, 2025; Available online: September 15, 2025)

Abstract

Accurate forecasting of Electric Vehicle (EV) sales is essential for supporting strategic decisions by policymakers, manufacturers, and investors amid the global shift toward sustainable transportation. This study compares the performance of two time series models, AutoRegressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) using historical EV sales data from 2010 to 2023. The ARIMA model, which is suited for linear trend projection, forecasts continued exponential growth, estimating sales to surpass 103 million units by 2025. In contrast, the LSTM model, known for capturing non-linear and complex patterns, projects a more moderate trend, with sales peaking at around 11.5 million units in 2022 before gradually declining. Evaluation using Mean Squared Error (MSE) shows that LSTM significantly outperforms ARIMA, achieving a lower error value (2.23×10^{14} vs. 4.44×10^{15}), indicating superior predictive accuracy. These results suggest that while ARIMA may be effective for short-term forecasting in stable markets, it can lead to overestimations in more dynamic environments. LSTM, with its ability to learn complex temporal dependencies, presents a more flexible and realistic tool for long-term planning in the evolving EV sector. The study contributes methodologically by offering a comparative analysis of two popular forecasting techniques and practically by guiding stakeholders on model selection. However, it is limited by its reliance on historical data and exclusion of external variables such as energy prices or policy changes. Future work should incorporate hybrid models and multi-source data to enhance forecasting robustness in the fast-changing EV market.

Keywords: ARIMA, Electric Vehicle, Forecasting, LSTM, Time Series

1. Introduction

The development of Electric Vehicles (EVs) is crucial for addressing global environmental challenges, particularly the reduction of greenhouse gas emissions. As part of the effort to mitigate climate change, EVs such as Battery Electric Vehicles (BEVs), Plug-in Hybrid Electric Vehicles (PHEVs), and Hybrid Electric Vehicles (HEVs) are pivotal in reducing tailpipe emissions and increasing the potential use of renewable energy sources [1], [2]. This transition aligns with international environmental commitments, such as the Paris Agreement, which aims to achieve net-zero emissions by mid-century [2], [3]. As more governments and industries focus on environmental sustainability, the role of EVs in reducing urban air pollution and supporting sustainability targets becomes even more significant. Studies have shown that financial incentives and the development of charging infrastructure are key to overcoming barriers and accelerating the adoption of EVs [4]. In this context, fostering a supportive policy environment and understanding consumer behavior are essential for achieving a faster and more widespread market penetration of EVs [4].

The global EV market is experiencing rapid growth, driven by government policies, technological advancements, and increasing environmental awareness. Forecasts suggest that EV sales could reach between 13 to 18 million units annually by 2025, with the potential to expand to 26 to 43 million units by 2030, capturing around 30% of the global automotive market share [5]. Governments worldwide are implementing policies that include subsidies, tax incentives, and regulations designed to reduce carbon emissions, which encourage consumers to adopt EVs [6]. Technological advancements, especially in battery technology, have further enhanced vehicle performance and boosted consumer confidence in EVs [7]. Alongside this, the growing availability of EV charging infrastructure has alleviated concerns

[*]**Corresponding author: Bhavana Srinivasan (bs@myjainuniversity.com)**
DOI: https://doi.org/10.47738/ijaim.v5i3.106

ISSN 2776-8007

about range anxiety, further supporting the transition [8]. Additionally, rising environmental awareness has led consumers to increasingly prioritize sustainable transportation options, emphasizing EVs as a cleaner, greener choice [6]. These intertwined factors have set the stage for continued expansion of the EV market, disrupting traditional automotive models and positioning EVs as a central player in future transportation solutions [9].

The importance of forecasting EV sales has become more apparent as it plays a critical role in enabling manufacturers, policymakers, and other stakeholders to plan long-term strategies effectively. Reliable forecasting can inform government regulations and help shape the financial incentives necessary to promote EV adoption, aligning with global environmental goals such as the Paris Agreement [10]. Furthermore, accurate predictions assist manufacturers in optimizing production schedules and supply chain management to meet market demand, avoiding inefficiencies that could lead to overproduction or shortages [11]. Forecasting models, such as Principal Component Analysis combined with General Regression Neural Networks, have proven useful in generating data-driven insights to guide decision-making processes across various sectors [12]. Understanding consumer behavior, particularly the growing emphasis on sustainability, can further aid stakeholders in adjusting marketing strategies and product offerings to align with consumer preferences [13]. Ultimately, integrating forecasting tools supports the development of robust charging infrastructure and networks, which are essential for maintaining momentum in the EV market and addressing range anxiety [14].

However, predicting the demand and sales of EVs remains a challenging task, mainly due to the influence of numerous external factors. The rapidly evolving nature of the EV market, driven by technological advancements, fluctuating energy prices, and changing consumer preferences, complicates the forecasting process [15]. New technologies often introduce unforeseen shifts in market dynamics, leading to unpredictable changes in adoption rates [16]. Additionally, global events, such as the COVID-19 pandemic, have significantly disrupted supply chains, reduced consumer purchasing power, and altered transportation needs, resulting in unpredictable effects on EV sales [2]. Moreover, government policies play a dual role; while they can promote EV adoption through incentives, sudden regulatory changes can create market uncertainty [17]. Incorporating these diverse factors, including the use of machine learning techniques, offers a promising solution to improve forecasting accuracy. Nonetheless, the inherent complexities of the EV market still pose significant barriers to achieving reliable predictions [16].

Given the unpredictable nature of the EV market and the reliance on outdated or overly simplistic models, the lack of accurate forecasting poses substantial risks to both business and governmental decision-making. Inaccurate forecasts can result in inefficient production schedules, inventory surplus, or shortages, which may hurt manufacturers' ability to meet market demand [18]. Moreover, relying on outdated models can lead to misguided government policies related to subsidies, infrastructure development, and regulations, hindering the growth of the EV market [18]. As the sector continues to evolve, integrating advanced analytics and machine learning algorithms is crucial for enhancing forecasting accuracy, enabling more informed decision-making across the board [19]. Without reliable forecasting models, strategic planning for EV adoption and environmental goals may be compromised, underscoring the need for sophisticated forecasting models to synchronize supply and demand in this rapidly changing market [15].

In conclusion, the need for more precise forecasting models to predict EV sales trends is vital due to the significant fluctuations observed in the market. Current methods often fail to capture the complex dynamics influencing EV sales, including technological advancements, governmental policies, and consumer behavior. Advanced models that combine historical data with analytics methodologies, such as General Regression Neural Networks (GRNN) and Principal Component Analysis (PCA), can offer more accurate sales predictions [12]. These models allow for better planning in terms of infrastructure, resource allocation, and marketing strategies, ultimately supporting the sustainable growth of the EV market. By incorporating socio-economic factors and market dynamics, forecasting models can be refined to provide more precise and reliable projections, which is essential for long-term success in the EV industry [2], [20].

2. Literature Review

2.1. The Role of Electric Vehicles in Sustainable Transportation

EVs have become a cornerstone in the effort to transition towards sustainable transportation. As global concerns over climate change and environmental degradation increase, the shift from traditional fossil fuel-based vehicles to electric

ISSN 2776-8007

alternatives has gained significant attention. The primary motivation behind the development and adoption of EVs is their potential to reduce greenhouse gas emissions, particularly Carbon Dioxide (CO2), which are major contributors to global warming. By using electricity as a power source, EVs generate zero tailpipe emissions, helping to improve urban air quality and reduce pollution levels in cities. This makes them an essential component of efforts to address urban pollution and promote cleaner environments [21].

Beyond their environmental benefits, EVs also offer energy efficiency advantages over conventional Internal Combustion Engine (ICE) vehicles. EVs convert a larger portion of the energy from their batteries into motion, making them more efficient compared to gasoline and diesel vehicles, which waste a considerable amount of energy as heat [22]. As battery technology improves, the operational costs of EVs, including maintenance and fuel, continue to decrease, making them a more affordable option for consumers in the long run [23]. The growing consumer interest in EVs is further supported by the shift towards renewable energy sources, which provides a cleaner way of charging EVs, reducing dependency on fossil fuels [24].

Despite these advantages, the widespread adoption of EVs is still hindered by several barriers, including the availability of charging infrastructure, the high cost of batteries, and varying government policies across regions. Investments in public charging infrastructure and financial incentives are critical to overcoming these barriers and encouraging broader adoption [25]. As the EV market evolves, overcoming these challenges is essential for realizing the full potential of EVs in reducing carbon emissions and fostering sustainable transportation.

2.2. Forecasting EV Sales and Market Trends

The growth of the EV market has been exponential in recent years, and accurate forecasting of future EV sales is becoming increasingly crucial for various stakeholders, including manufacturers, policymakers, and investors. Predicting sales trends allows manufacturers to adjust production schedules, policymakers to plan for necessary infrastructure development, and investors to allocate resources effectively. Among the various forecasting models, ARIMA and LSTM are two of the most widely used approaches for time series forecasting.

ARIMA is a statistical model that is particularly effective for forecasting stationary time series data. It captures temporal dependencies by combining Autoregressive (AR), Integrated (I), and Moving Average (MA) components. ARIMA works well for forecasting sales trends when the data exhibits consistent patterns or trends over time, making it a popular choice for automotive sales predictions. Studies have shown its effectiveness in predicting EV sales, especially when analyzing the impact of external factors such as socio-political events or market shocks. For example, Wu [18] applied ARIMA to analyze the impact of the Russia-Ukraine conflict on Tesla's sales, revealing how geopolitical events can disrupt market behavior. However, ARIMA models often require stationary data, meaning that the data must have a constant mean and variance over time, which may not always be the case in the dynamic EV market.

In contrast, LSTM, a type of Recurrent Neural Network (RNN), excels at modeling non-linear relationships and capturing long-term dependencies within sequential data. LSTM is particularly useful for forecasting EV sales in rapidly evolving markets, where consumer preferences and technological innovations continuously change. LSTM can learn complex temporal patterns, making it more suitable for forecasting in industries like electric vehicles, where market dynamics are affected by both internal and external factors. As consumer preferences shift and new technologies emerge, LSTM provides more flexibility and can adapt to these changes better than traditional models like ARIMA. Its ability to predict long-term trends based on historical data makes it highly effective for dynamic sectors such as EV sales forecasting, where rapid market fluctuations are common.

2.3. Challenges and Future Directions in EV Sales Forecasting

Forecasting the sales of EVs involves numerous challenges, as the market is influenced by a wide range of dynamic factors, including governmental policies, technological advancements, energy prices, and consumer behavior. One of the primary challenges is the inconsistency and volatility of government policies, which can create uncertainty in the EV market. Inconsistent subsidies or shifting regulations can disrupt market behavior, making it difficult for both consumers and manufacturers to predict future trends. For example, fluctuating incentives or sudden policy changes regarding emissions standards can alter consumer purchase intentions, making forecasting efforts more complex [26].

In regions where EV adoption is incentivized by financial rewards, any reduction in these incentives can cause a temporary decline in demand, leading to inaccurate predictions.

Another challenge is the impact of energy prices on EV adoption. While EVs are often marketed as a cleaner and more cost-effective alternative to traditional vehicles, their operating costs are highly sensitive to energy prices. In regions where electricity prices fluctuate significantly, consumers may hesitate to switch to electric vehicles, especially if their perceived long-term savings are undermined by volatile energy costs. As energy prices affect consumer behavior, they must be integrated into forecasting models to better reflect market realities [27].

The availability of charging infrastructure also plays a crucial role in determining EV sales. Research has shown that a lack of sufficient charging stations can deter potential EV buyers due to concerns over range anxiety and the convenience of finding charging stations. This is particularly true in developing markets where infrastructure development may lag behind the demand for EVs. The absence of a robust charging network not only affects consumer confidence but also hinders manufacturers' ability to scale production in line with demand. Thus, effective forecasting requires incorporating the state of charging infrastructure into models, alongside technological and policy changes, to provide more accurate predictions [28].

Despite these challenges, the integration of machine learning techniques and hybrid forecasting models offers promising directions for improving forecasting accuracy. For example, combining traditional statistical models like ARIMA with machine learning approaches such as LSTM can enhance predictive capabilities by capturing both linear and non-linear relationships in the data. Additionally, incorporating real-time data, such as consumer sentiment or online purchasing trends, can further refine forecasts by reflecting market dynamics as they evolve [29]. As the EV market continues to grow and evolve, it will be essential for future forecasting models to adapt to these complexities and integrate external variables more effectively to maintain their relevance and accuracy.

3. Methodology

To provide a clear overview of the steps involved in the forecasting process, a structured workflow is presented in figure 1. This figure illustrates the complete pipeline, starting from data collection, followed by comprehensive preprocessing steps, including data cleaning, transformation, and time series structuring. It then continues with the application of forecasting techniques such as ARIMA and LSTM, and concludes with model evaluation using MSE. The diagram serves as a visual guide to understand how raw data is transformed into actionable forecasting insights.

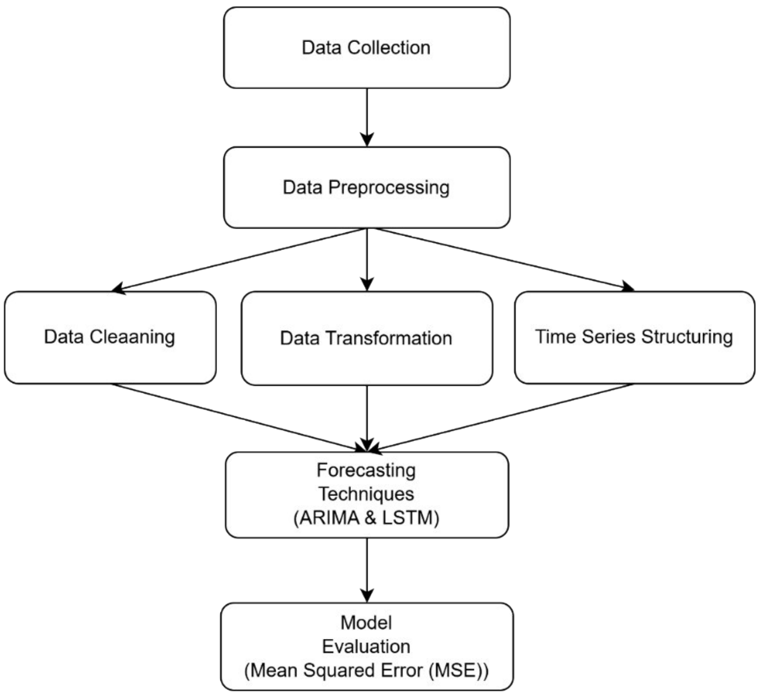

Figure 1. Research Methodology

ISSN 2776-8007

3.1. Data Collection

The dataset used in this study provides historical data on EV sales, structured in several key columns. The information is captured across various regions, vehicle categories, and sales parameters. The dataset is designed to track the growth and adoption of electric vehicles over time, providing valuable insights for time series forecasting. Table 1 provides a detailed overview of the dataset's structure:

Table 1. Dataset Overview

Column Name	Description
Region	Geographical region where the EV sales are tracked (e.g., specific countries, continents, or regions).
Category	Classification of the data (e.g., "Historical" refers to actual recorded data of EV sales).
Parameter	Key variable being measured (e.g., "EV sales" refers to total number of electric vehicles sold).
Mode	Vehicle type category (e.g., "Cars," "Light Commercial Vehicles").
Powertrain	Type of technology used in the vehicle (e.g., "BEV" for Battery Electric Vehicle, "PHEV" for Plug-in Hybrid Electric Vehicle).
Unit	Measurement unit for the parameter (e.g., "Vehicles" for number of vehicles, "Percent" for market share).
Year	Year for the data point (e.g., 2011, 2012, etc.).
Value	The actual value being reported for the given parameter (e.g., the number of vehicles sold or market share in percentage).

The dataset used in this study is obtained from Kaggle, a popular platform for data science and machine learning competitions. The dataset includes historical data on EV sales, sourced from various global reports and industry sources. It provides insights into EV adoption trends, sales figures, and market growth patterns. By using this Kaggle dataset, we aim to build a time series forecasting model to predict future EV sales growth based on historical data.

3.2. Data Preprocessing

Before performing the analysis, several data cleaning and preprocessing steps were carried out to ensure the dataset was ready for forecasting. First, missing values were identified and handled by either removing the rows with missing data or imputing values based on available data using techniques like mean imputation or forward filling. We also focused on filtering relevant data, keeping only the columns related to electric vehicle sales, such as region, category, parameter, year, and value, and removing any unrelated data. The data types for each column were verified and corrected; for example, the year column was ensured to be in integer format, and the value column was set to a numeric type to enable numerical operations.

Additionally, we addressed outliers that could distort the forecasting process, either by removing or adjusting extreme values that were inconsistent with the rest of the dataset. For models like LSTM, which are sensitive to feature scaling, we applied MinMaxScaler to normalize the data, ensuring all values were within the same scale and improving model performance. Finally, the dataset was structured in a time series format, with the year as the index and sales data as the values, making it suitable for time-dependent analysis and forecasting. These preprocessing steps ensured that the dataset was clean, consistent, and ready for effective analysis.

3.3. Forecasting Techniques

ARIMA is a popular statistical model for forecasting time series data. It combines three components. The AutoRegressive component models the relationship between an observation and a specified number of lagged observations. The Integrated component involves differencing the time series data to make it stationary, ensuring that the data has a constant mean and variance over time. The Moving Average component captures the relationship between an observation and the residual errors from a moving average model applied to previous time steps. The general form of the ARIMA model is represented as:

$$Y_t = \mu + \varnothing_1 Y_{t-1} + \varnothing_2 Y_{t-2} + \dots + \varnothing_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (1)$$

Note:
Y_t is the current value.
$\varnothing$ are the parameters of the AR term.

 ISSN 2776-8007

θ are the parameters of the MA term.
ϵ_t is the error term.
p, d, and q are the parameters for AR, differencing, and MA respectively.

ARIMA is particularly useful when data exhibits trends or seasonality, provided the series is made stationary through differencing. The model uses past values of the data and residuals to predict future values, making it effective for short-term forecasting, especially in situations where past observations are reliable indicators of future trends.

LSTM is a type of RNN that is specifically designed to handle sequential data and model long-term dependencies. LSTM networks consist of memory cells that store information for long periods, allowing the model to learn complex patterns and relationships over time. Unlike traditional RNNs, which struggle with long-term dependencies, LSTM's architecture is designed to retain useful information and discard irrelevant data. This makes LSTM particularly effective for time series forecasting, where past observations can influence future predictions in both short-term and long-term contexts.

The general equation for LSTM is represented by the following set of equations that control the flow of information within the LSTM units:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{5}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{7}$$

Note:
f_t is the forget gate that determines how much of the past information should be kept.
i_t is the input gate that controls how much new information is added to the cell state.
$\tilde{C}_t$ is the candidate cell state, which is a proposed update to the cell state.
C_t is the cell state, which stores information over time.
o_t is the output gate, determining what part of the cell state is output to the next hidden state.
h_t is the hidden state, which is used for output.

LSTM is particularly advantageous for predicting complex data patterns, such as those found in electric vehicle sales, where historical data may contain both long-term trends and short-term fluctuations. The model's ability to capture these dependencies makes it highly effective for time series forecasting in industries that experience rapid changes, like the EV market.

3.4. Model Evaluation

To evaluate the performance and accuracy of the forecasting models (ARIMA and LSTM), we use several commonly employed metrics in time series forecasting. One of the key metrics is the Mean Squared Error (MSE).

MSE is commonly used to quantify the accuracy of a model by measuring the average squared difference between the actual and predicted values. Unlike MAE, which uses the absolute differences, MSE gives more weight to larger errors due to the squaring of differences, making it more sensitive to outliers. The formula for MSE is:

$$MSE = \frac{1}{n}\sum_{t-1}^{n}(\mathscr{Y}_t - \hat{\mathscr{Y}}_t)^2 \tag{8}$$

Note:
$\mathscr{Y}_t$ is the actual value at time t,

 ISSN 2776-8007

$\hat{\mathscr{Y}}_t$ is the predicted value at time t,
n is the number of data points in the test set.
In this formula:
$(\mathscr{Y}_t - \hat{\mathscr{Y}}_t)^2$ represents the squared error at each time step, penalizing larger errors more heavily.

The sum of these squared errors is averaged over all data points in the test set. Lower MSE values indicate better performance, meaning the model's predictions are closer to the actual values. Higher MSE values indicate larger discrepancies between predicted and actual values, signaling that the model is less accurate.

MSE is particularly useful when larger errors are more problematic, as it penalizes such discrepancies more than MAE. This makes MSE suitable for cases where larger deviations from the true values are more critical to the forecasting task.

For both the ARIMA and LSTM models, the MSE is calculated by comparing the predicted values with the actual data points, allowing us to evaluate the overall forecasting accuracy and determine which model provides the best fit for the EV sales data.

4. Results and Discussion

4.1. Result

Figure 2 presents the historical curve of EV sales from 2010 to 2023. This graph provides a comprehensive overview of how the EV industry has grown over the past decade. It is evident that during the early period (2010–2017), growth occurred gradually. However, starting around 2018 through 2023, a significant surge in EV adoption can be observed, indicating accelerating market acceptance.

This sharp increase can be attributed to several external factors, including government incentives, advancements in battery technology, and growing environmental awareness. The exponential trend seen during this period serves as a rational foundation for applying time series forecasting models like ARIMA, which are inherently suited to capturing long-term linear trends based on historical patterns.

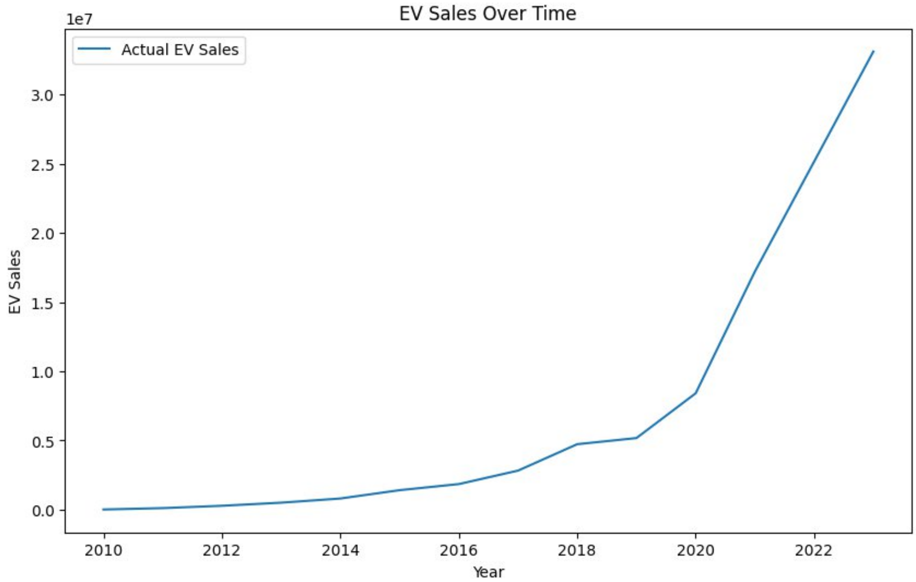

Figure 2. EV Sales Over Time

By analyzing the actual data trends visualized in figure 2, we gain a strong justification for modeling EV sales predictions. It demonstrates that the historical data contains a distinct growth pattern, making it viable to project future values statistically. Therefore, the use of the ARIMA model in the following section aims to estimate the continuation of this trend into the near future.

 ISSN 2776-8007

Figure 3 illustrates the ARIMA model's forecast for EV sales from 2021 to 2025, derived from the historical trend previously shown in figure 2. In this graph, the solid blue line represents actual EV sales data from 2010 to 2023, while the dashed orange line shows the projected values generated by the ARIMA model. The visual demonstrates a continued growth trajectory that extends from the historical data into the forecasted period.

According to the ARIMA forecast, EV sales are expected to rise sharply, increasing from approximately 46 million units in 2021 to over 103 million units by 2025. This exponential projection reflects ARIMA's reliance on historical momentum, where past upward trends are extended mathematically into the future. While such forecasts can be useful for short-term planning in fast-growing markets, they may overlook external constraints such as market saturation, policy changes, or supply chain issues. Therefore, ARIMA projections should be interpreted with caution, especially when used for long-term strategic decision-making.

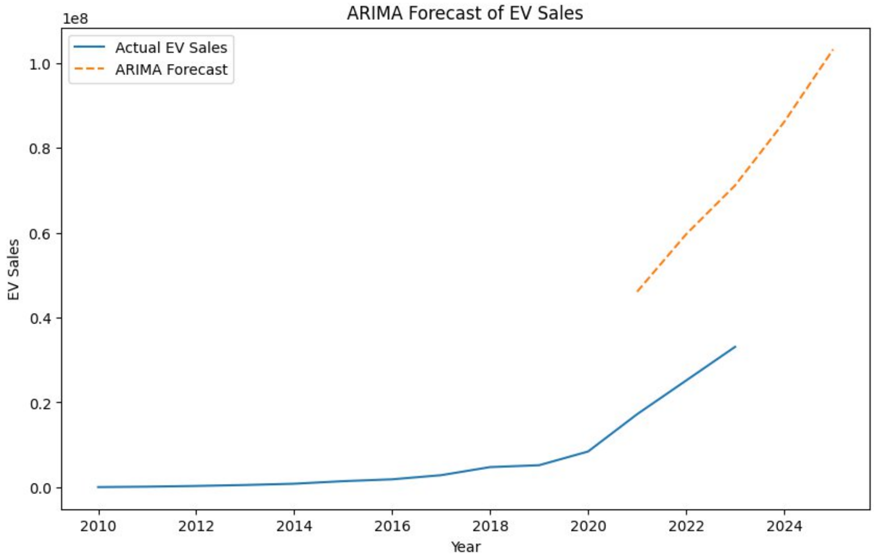

Figure 3. ARIMA Forecast vs Actual EV Sales

Table 2. ARIMA Predictions

Year	ARIMA Predictions (Vehicles)
2021	46,091,350
2022	59,636,640
2023	71,158,170
2024	86,107,010
2025	103,225,500

Table 2 presents the numerical results of the ARIMA model's forecast for EV sales from 2021 to 2025. The values confirm the steep upward trend depicted in figure 3. Starting from an estimated 46 million units in 2021, ARIMA forecasts a dramatic increase to approximately 103 million units by 2025.

This consistent year-over-year growth underlines ARIMA's nature as a time series model that projects forward based on past momentum. While such projections are mathematically coherent, the absolute values reinforce the need for cautious interpretation, as the model assumes uninterrupted exponential growth without adjusting for possible market saturation or external constraints.

To address the limitations of linear models like ARIMA, a LSTM neural network model was also applied. Figure 4 presents a visual comparison of both ARIMA and LSTM forecasts against actual EV sales data. In this figure, the solid blue line shows the historical EV sales, the dashed orange line represents the ARIMA forecast (as seen previously in figure 3), and the dashed green line depicts the LSTM forecast. Unlike ARIMA, which continues the steep upward trend, the LSTM model predicts a more moderate and plateauing trajectory. It estimates that EV sales will peak around 11.5 million units in 2022, followed by a gradual decline to approximately 7 million units by 2025.

 ISSN 2776-8007

This contrast reveals how the choice of forecasting model can shape strategic interpretation. While ARIMA tends to extend existing trends and assumes uninterrupted growth, LSTM incorporates more nuanced temporal patterns and potential non-linear influences. As a result, LSTM offers a more cautious and arguably more realistic forecast, making it a valuable tool for decision-makers seeking long-term stability and risk awareness.

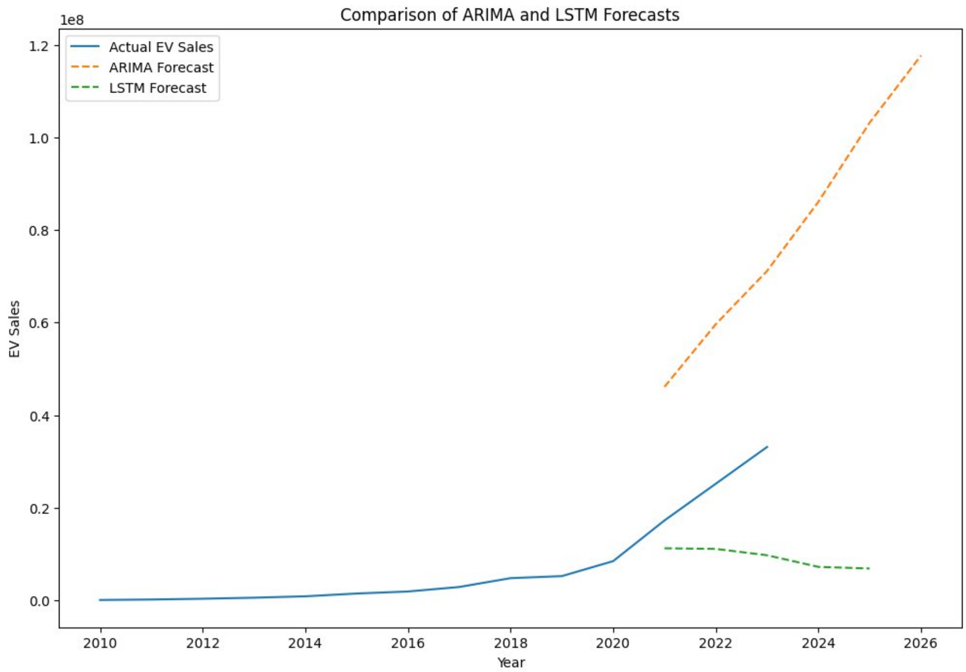

Figure 4. Comparison of ARIMA and LSTM Forecasts

Table 3. LSTM Predictions

Year	LSTM Predictions (Vehicles)
2021	11,432,678
2022	11,525,443
2023	10,104,892
2024	7,327,283
2025	6,960,020

Table 3 displays the forecasted EV sales generated by the LSTM model. Unlike ARIMA, the LSTM predictions reflect a more conservative and nuanced trajectory. According to this model, EV sales are expected to peak at around 11.5 million units in 2022, then decline gradually to approximately 7 million units by 2025.

These figures are consistent with the trend shown in figure 4, suggesting that the LSTM model is more sensitive to changes in temporal dynamics and non-linear influences. The plateau and decline captured in the table reinforce LSTM's capability to account for market slowdowns or saturation effects, making it a more realistic and cautious forecasting tool for long-term planning.

Table 4. MSE Comparison

Model	Mean Squared Error (MSE)
ARIMA	4.44×10^{15}
LSTM	2.23×10^{14}

To objectively assess the forecasting accuracy of the ARIMA and LSTM models, a comparison was conducted using the MSE metric, as presented in table 4. MSE measures the average squared difference between predicted and actual values, where a lower value indicates a model with better predictive accuracy.

The results clearly demonstrate that the LSTM model outperforms ARIMA, with a significantly lower MSE of 2.23×10^{14}, compared to ARIMA’s much higher value of 4.44×10^{15}. This substantial difference quantitatively supports the earlier visual and numerical analysis, which already suggested that ARIMA tends to overestimate future EV sales by extending past trends too aggressively. In contrast, the LSTM model captures market dynamics more realistically by incorporating non-linear patterns and temporal dependencies.

These findings highlight the importance of model selection in time series forecasting. While ARIMA can be useful for short-term trend continuation, LSTM offers a more reliable foundation for long-term planning, particularly in rapidly evolving markets like electric vehicles. The superior performance of LSTM, as shown through its lower MSE, underscores its suitability for scenarios requiring stable, accurate, and context-sensitive forecasts.

4.2. Discussion

The results of this study reaffirm the rapid and transformative growth of the EV market over the past decade, as reflected in figure 2. The exponential increase in EV adoption from 2018 to 2023 aligns with prior findings that attribute this trend to supportive government policies, advancements in battery technology, and increasing environmental awareness [21]. These factors have positioned EVs as central to the global strategy for sustainable transportation, especially given their superior energy efficiency and potential to reduce urban emissions [22], [23].

The ARIMA model's projection in figure 3 and table 2 extends this exponential trend, forecasting EV sales to more than double between 2021 and 2025. This reflects ARIMA's statistical strength in capturing and extrapolating linear trends in time series data, as also demonstrated in previous studies such as Wu [18], who used ARIMA to assess the impact of geopolitical events on EV markets. However, as this study and others have pointed out, ARIMA assumes stationarity and consistent variance, making it less responsive to the complex, rapidly shifting conditions of the EV landscape [26], [27]. As such, while ARIMA offers valuable short-term insights, it may overestimate long-term sales if it fails to account for market saturation, infrastructure limitations, or policy volatility.

In contrast, the LSTM model presents a more conservative outlook, as seen in figure 4 and table 3, with sales peaking around 11.5 million units in 2022 before gradually declining. This aligns with LSTM's capability to handle non-linear, dynamic relationships in time series data, making it well-suited for markets where external factors such as policy changes, infrastructure development, and consumer behavior play a critical role [24]. By recognizing pattern shifts and temporal dependencies, LSTM offers a more adaptable approach for long-term forecasting in evolving markets like EVs, consistent with literature emphasizing machine learning's flexibility over traditional models [29].

From a methodological standpoint, this study contributes to the existing literature by providing a direct comparative evaluation of ARIMA and LSTM models using real-world EV sales data. While many prior studies have focused on a single approach, the dual-model comparison presented here highlights how model selection impacts both the predicted trends and the strategic conclusions derived from them. The integration of forecast visualization, numerical output, and performance metrics (e.g., MSE) adds depth to the analysis and offers a replicable framework for future forecasting in other emerging technology markets.

The implications of this research are particularly relevant for policymakers, manufacturers, and investors. For instance, overreliance on optimistic forecasts such as those produced by ARIMA may lead to overproduction or misallocation of resources. Conversely, more grounded forecasts like those from LSTM may support more sustainable infrastructure development and realistic policy targets. Understanding these model differences enables better planning in areas such as battery supply chains, charging infrastructure, and emission regulation strategies.

However, this study is not without limitations. First, both models rely solely on historical sales data and do not directly integrate external influencing variables such as fuel prices, regulation shifts, or consumer sentiment, which are known to affect EV market behavior [27], [28]. Second, while LSTM is powerful in handling sequential data, it is also sensitive to hyperparameter tuning and training data quality, which may limit its generalizability if not properly managed. Finally, the projection period (2021–2025) is relatively short; longer forecasting horizons may require hybrid or ensemble approaches that combine statistical and machine learning techniques.

Despite these limitations, the findings underline the critical importance of selecting forecasting models that align with the specific dynamics of the market in question. As the EV industry continues to evolve under the influence of technological, political, and economic pressures, adaptive and data-rich models like LSTM are likely to become increasingly essential for forward-looking decision-making.

ISSN 2776-8007

5. Conclusion

This study aimed to forecast EV sales using two time series models, ARIMA and LSTM based on historical sales data. The analysis confirmed that EV adoption has experienced exponential growth in recent years, driven by government incentives, technological advancements, and rising environmental awareness. The ARIMA model projected a continuous, steep increase in sales, reflecting a linear growth assumption rooted in historical momentum. In contrast, the LSTM model offered a more conservative forecast, anticipating a sales peak around 2022 followed by a gradual decline. These contrasting results underscore the importance of model selection, as each approach influences how trends are interpreted and how strategies are formed. While ARIMA is useful for short-term forecasts in stable conditions, it lacks flexibility in responding to non-linear or disruptive market factors. LSTM, with its ability to model complex temporal dynamics, presents a more realistic and adaptive approach for forecasting in volatile or rapidly evolving markets.

Quantitative evaluation using the MSE metric further validated the models' performances, revealing that LSTM achieved significantly lower error values than ARIMA, thereby confirming its superior predictive accuracy. This study contributes to the literature by offering a comparative evaluation of traditional and deep learning models in the context of EV forecasting, with direct implications for policymakers, manufacturers, and investors. However, its reliance solely on historical sales data limits the depth of analysis, as external variables like policy changes, energy prices, or infrastructure development were not included. Future studies should explore hybrid models that integrate statistical and machine learning techniques, while incorporating real-time and multi-dimensional data to enhance forecast reliability. In conclusion, LSTM emerges as a more effective forecasting tool for anticipating EV market trends and informing data-driven decision-making in the transportation sector.

6. Declarations

6.1. Author Contributions

Conceptualization: B.S.; Methodology: B.S.; Software: B.S.; Validation: B.S.; Formal Analysis: B.S.; Investigation: B.S.; Resources: B.S.; Data Curation: B.S.; Writing – Original Draft Preparation: B.S.; Writing – Review and Editing: B.S.; Visualization: B.S.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing 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

[1] X. Long, C. Liu, B. Miao, J. Xiao, Y. Cao, Z. Xu, C. Li, L. Wang, and Q. Yang, "Life Cycle Assessment of Typical New Energy Vehicles in China," *J. Phys. Conf. Ser.*, vol. 3000, no. 1, pp. 1–9, 2025, doi: 10.1088/1742-6596/3000/1/012016.

[2] E. Javanmardi, M. Hoque, A. Tauheed, and M. Umar, "Evaluating the Factors Affecting Electric Vehicles Adoption Considering the Sustainable Development Level," *World Electr. Veh. J.*, vol. 14, no. 5, pp. 1–19, 2023, doi: 10.3390/wevj14050120.

[3] D. Nasreddin, H. E. Hafdaoui, F. Jelti, A. Boumelha, and A. Khallaayoun, "Inhibitors of Battery Electric Vehicle Adoption in Morocco," *World Electr. Veh. J.*, vol. 15, no. 1, pp. 6, 2023, doi: 10.3390/wevj15010006.

[4] Q. Jiaqi, Y. Zhang, and C. Fengming, "Analyzing the Determinants Affecting Chinese Consumers Willingness to Adopt Electric Vehicles," *Acad. J. Bus. Manag.*, vol. 5, no. 22, pp. 208–216, 2023, doi: 10.25236/ajbm.2023.052230.

[5] T. Wiśniewski and B. Tundys, "Determinants of Growth of the Electric Car Market – Investigating on the Truck Market," *Eur. Res. Stud. J.*, vol. XXV, no. 2, pp. 210–222, 2022, doi: 10.35808/ersj/2916.

[6] H. N. D. Senyapar, M. Akil, and E. Dokur, "Adoption of Electric Vehicles: Purchase Intentions and Consumer Behaviors Research in Turkey," *Sage Open*, vol. 13, no. 2, pp. 1–17, 2023, doi: 10.1177/21582440231180586.

[7] O. A. Marzouk, "Growth in the Worldwide Stock of E-Mobility Vehicles (By Technology and by Transport Mode) and the Worldwide Stock of Hydrogen Refueling Stations and Electric Charging Points Between 2020 and 2022," *Key Eng. Mater.*, vol. 972, no. December, pp. 89–96, 2023, doi: 10.4028/p-8imgm4.

[8] E. A. Etukudoh, A. Hamdan, V. I. Ilojianya, C. D. Daudu, and A. Fabuyide, "Electric Vehicle Charging Infrastructure: A Comparative Review in Canada, Usa, and Africa," *Eng. Sci. Technol. J.*, vol. 5, no. 1, pp. 245–258, 2024, doi: 10.51594/estj.v5i1.747.

[9] C. Hecht, K. G. Spreuer, J. Figgener, and D. U. Sauer, "Market Review and Technical Properties of Electric Vehicles in Germany," *Vehicles*, vol. 4, no. 4, pp. 903–916, 2022, doi: 10.3390/vehicles4040049.

[10] R. Cai, L. Li, and W. Lei, "Are Residents Willing to Pay for Electric Cars? An Evolutionary Game Analysis of Electric Vehicle Promotion in Macao," *World Electr. Veh. J.*, vol. 15, no. 8, pp. 1–20, 2024, doi: 10.3390/wevj15080371.

[11] Z. Min, Y. Fang, and D. Dai, "Forecast of the Evolution Trend of Total Vehicle Sales and Power Structure of China Under Different Scenarios," *Sustainability*, vol. 15, no. 5, pp. 1–22, 2023, doi: 10.3390/su15053985.

[12] M. Wu and W. Chen, "Forecast of Electric Vehicle Sales in the World and China Based on PCA-GRNN," *Sustainability*, vol. 14, no. 4, pp. 1–14, 2022, doi: 10.3390/su14042206.

[13] S. C. Arora and V. K. Singh, "Transition to Green Mobility: A Twin Investigation Behind the Purchase Reasons of Electric Vehicles in the Indian Market," *Bottom Line Manag. Libr. Financ.*, vol. 37, no. 3, pp. 277–308, 2024, doi: 10.1108/bl-08-2023-0249.

[14] T. Aldhanhani, A. Abraham, W. Hamidouche, and M. F. Shaaban, "Future Trends in Smart Green IoV: Vehicle-to-Everything in the Era of Electric Vehicles," *IEEE Open J. Veh. Technol.*, vol. 5, no. January, pp. 278–297, 2024, doi: 10.1109/ojvt.2024.3358893.

[15] Z. Han, D. Zhao, F. Sun, and H. Zhu, "A Study on the Diffusion Model of New Energy Passenger Vehicles with Consideration of Product Value," *PLoS One*, vol. 20, no. 5, pp. e0323316, 2025, doi: 10.1371/journal.pone.0323316.

[16] F. Teng, "Beyond Traditional Forecasting: Machine Learning and Adaptive Algorithms in EV Sales Predictions," *Appl. Comput. Eng.*, vol. 82, no. 1, pp. 12–16, 2024, doi: 10.54254/2755-2721/82/2024glg0081.

[17] J. Chen, Z. Zhang, C. Zhao, S. Zhang, W. Guo, C. Lu, and X. Sun, "Research on the Purchase Intention of Electric Vehicles Based on Customer Evaluation and Personal Information," *World Electr. Veh. J.*, vol. 15, no. 1, pp. 1–17, 2023, doi: 10.3390/wevj15010009.

[18] S. Wu, "Predicting Stock Prices in the Electric Vehicle and Technology Sectors Using Long Short-Term Memory Models," *Adv. Econ. Manag. Polit. Sci.*, vol. 94, no. 1, pp. 148–156, 2024, doi: 10.54254/2754-1169/94/2024ox0183.

[19] Z. Wei, "Forecasting Development Trends and Analysing Influencing Factors of New Energy Electric Vehicles in China: Based on Multi-Level Analysis and Time Series Modelling," *Math. Model. Algorithm Appl.*, vol. 2, no. 1, pp. 44–48, 2024, doi: 10.54097/1pydn504.

[20] O. Egbue, S. Long, and S. D. Kim, "Resource Availability and Implications for the Development of Plug-in Electric Vehicles," *Sustainability*, vol. 14, no. 3, pp. 1–17, 2022, doi: 10.3390/su14031665.

[21] A. Lazuardy, "Evaluation of Electric Car Research Trends Through Bibliometric Analysis: A Literature Review," *Int. J. Sci. Technol.*, vol. 2, no. 3, pp. 108–116, 2023, doi: 10.56127/ijst.v2i3.1166.

[22] S. K. Mittal, P. Korde, S. Palaniraja, N. Omase, P. K. Guchhait, and P. Mundra, "A Systematic Review on Recent Advancement in Electric Vehicle Technologies," *Int. Res. J. Eng. Appl. Sci.*, vol. 11, no. 4, pp. 37–44, 2023, doi: 10.55083/irjeas.2023.v11i04006.

[23] A. Mousaei, Y. Naderi, and I. S. Bayram, "Advancing State of Charge Management in Electric Vehicles with Machine Learning: A Technological Review," *IEEE Access*, vol. 12, no. March, pp. 43255–43283, 2024, doi: 10.1109/access.2024.3378527.

[24] M. Ntombela, K. Musasa, and K. Moloi, "A Comprehensive Review for Battery Electric Vehicles (BEV) Drive Circuits Technology, Operations, and Challenges," *World Electr. Veh. J.*, vol. 14, no. 7, pp. 1–23, 2023, doi: 10.3390/wevj14070195.

[25] I. Ali and M. Naushad, "Insights on Electric Vehicle Adoption: Does Attitude Play a Mediating Role?," *Innov. Mark.*, vol. 18, no. 1, pp. 104–116, 2022, doi: 10.21511/im.18(1).2022.09.

[26] S. Hardman, A. Chakraborty, K. Hoogland, C. Sugihara, J. P. Helveston, A. F. Jensen, A. Jenn, P. Jochem, P. Plotz, F. Sprei, B. Williams, J. Axsen, E. Figenbaum, J. Pontes, G. Tal, and N. Refa, "Demand-Side Challenges and Research Needs on the Road to 100% Zero-Emission Vehicle Sales," *Prog. Energy*, vol. 7, no. 2, pp. 1–16, 2025, doi: 10.1088/2516-1083/ad9f1a.

[27] P. Gershon, D. Rempel, J. Mason, G. A. Macht, R. A. Arbelaez, and L. N. Boyle, "Advances and Challenges in Vehicle Electrification," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 68, no. 1, pp. 875–878, 2024, doi: 10.1177/10711813241261283.

[28] N. A. Q. Muzir, M. R. H. Mojumder, M. Hasanuzzaman, and J. Selvaraj, "Challenges of Electric Vehicles and Their Prospects in Malaysia: A Comprehensive Review," *Sustainability*, vol. 14, no. 14, pp. 1–40, 2022, doi: 10.3390/su14148320.

[29] Z. Yan, R. Zhao, and B. Liu, "Electric Vehicle Industry in Three Main RegionsUnited States, China, and EU Countries," *Adv. Econ. Manag. Polit. Sci.*, vol. 17, no. 1, pp. 181–185, 2023, doi: 10.54254/2754-1169/17/20231084.