E3S Web of Conferences **729**, 04003 (2026) https://doi.org/10.1051/e3sconf/202672904003
SPETM 2026

A comparative analysis of regression models and country-level adoption trajectories of electric vehicles

Olugbenga Falodun , *Fadesakin* Omowole*, *Samson Ojo* Ayileka , *Fasasi* Bello , *Adeniyi Samuel* Ige , and *Oladeji Lawrence* Ajilu

Department of Electrical and Electronics Engineering Technology, Rufus Giwa Polytechnic, Owo, Nigeria

Abstract. With an increase in electric vehicle (EV) adoption at a rapid pace, the need for quality forecasting models to establish different adoption patterns across multiple countries will accommodate for this need moving forward. This work provides analysis between seven regression algorithms for forecasting EV sales and looks to establish the country-level patterns of EV adoption using various clustering techniques. Historical data from the International Energy Agency (IEA) Global EV dataset was used to capture this data at the country level from 2010 through to 2024 across 43 countries worldwide. These historical data were used to perform predictive modelling using the following models; Linear Regression, Ridge Regression, Random Forest, Gradient Boosting, XGBoost, K-Nearest Neighbours (KNN), and Support Vector Regression (SVR), to track statistical performance of each model with respect to predictive modelling through Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and a Coefficient of Determination (R^2). In addition, Principal Component Analysis (PCA) and K-Means Clustering were used to identify and track country-level EV adoption trajectories. Of the models used to predict EV sales, Ridge Regression showed the greatest level of forecasting performance with the lowest levels of MAE, RMSE and the least negative R^2; while complex ensemble and kernel-based models showed high levels of overfitting as well as poor levels of generalization. The PCA results presented four country-level clusters and highlighted mature markets, high growth emerging economies and low volume adopters of EVs. Overall, the results indicate that using regularised linear models will yield more reliable short-term forecasting than more complex machine learning algorithms.

1 Introduction

Electric mobility's transformation across the globe has progressed rapidly over the last ten years, largely due to the pressure of climate change targets, advances in technological innovation for batteries, and favourable policy environments in numerous economies around the world [1]. Electric vehicles (EVs) are widely considered to be a keystone of sustainable transportation; however, there is considerable variance in adoption rates depending on geographic location. To allocate resources wisely and develop suitable interventions on time, policymakers, automakers, energy providers, and investors need to understand historical sales patterns, determine which markets are leading and lagging in Electric vehicle (EV) sales and project future EV sales accurately [2].

Increasingly, researchers are trying to build models and forecasts on how many EVs will be adopted. Most of these early studies used econometric/statistical methods, which represent 54% of all known forecasts published to date [3]. The vast majority of the early models used conventional time-series techniques to forecast EVs, including ARIMA and exponential smoothing, which have fairly well performed in the rate of growth of EVs when they are close to linear [4]. Although EV adoption databases can be treated as conventional time series, the fact that the adoption of EVs does not occur in a linear manner, requires a change in the forecasting methodology used to model the different growth cycles associated with the adoption of EVs. A rigorous literature review of 194 peer-reviewed studies conducted between 2016 and 2025 demonstrated that machine learning (ML) methods are being adopted as the primary methodology for forecasting EVs sales, e.g., through the combination of ensemble (18%), behavioural simulation (14%), hybrid models (8%), and deep learning (4%) techniques [3].

Studies have compared the performance of algorithms for EV prediction tasks within the machine learning framework. Study found that multiple machine learning methods consistently outperformed linear methods when forecasting EV adoption in the US using tree-based models [5]. In contrast, study evaluated the accuracy of ARIMA and LSTM models for global EV sales prediction, and determined that deep learning methods produced better long-term results but needed extensive hyperparameter tuning [2]. Ensemble models, including Random Forest, LightGBM, and XGBoost, provided superior forecasting results than traditional panel regression models across OECD countries between 2010 and 2023, with significant advantages in

* Corresponding author: fadesakinomowole@gmail.com

 https://doi.org/10.1051/e3sconf/202672904003

multidimensional policy contexts [6]. Contrary to these studies, ensemble models only marginally outperformed linear regression (R^2 = ~0.18 vs. ~0.08) when applied to an extensive global EV sales dataset, demonstrating that predicting EV sales trends is very complex [7].

With respect to the UK market, XGBoost was reported as the most accurate of the six different SARIMAX, Prophet, and XGBoost models; however, XGBoost's extrapolation capabilities were limited, thus requiring use of complementary time-series methods for longer-term forecasts [8].

The study presents a predictive modeling framework that utilizes seven regression algorithms, all of which have been developed based on established statistical learning theories as described in the research literature. Linear regression is the original parametric form of regression, where input features are combined in a linear fashion to create a weighted sum, and the goal is to minimize the sum of the squared residuals. It is the "benchmark" regression model because it is easy to interpret and the computation time is fast. However, because of the specific assumptions about linearity and homoscedasticity that must hold, it is not well suited to modeling the nonlinear aspects of EV adoption [9]. Ridge regression uses ordinary least squares as its starting point, but adds a penalty that shrinks the coefficients down to 0. This shrinkage reduces variance and aids in resolving multicollinearity issues and enhances the capacity to generalize when the predictors are correlated. Recent work has provided some insight on ridge regression in high-dimensional settings and the relationship between ridge regression and the coefficient of determination [10].

Although the previous research on EV sales forecasting shows considerable progress, there are still significant gaps in the literature. First, most published comparative studies look at only between two to four algorithms simultaneously; thus, performance rankings cannot necessarily be generalized across model families [5, 7]. Second, most of the research currently available focuses on contexts within a single country, without systematically exploring how models perform across heterogeneous geographic areas [6, 8]. As such, those who have studied the forecasts for one country may not have robustly informed cross-region inferences about the best models for forecasting EV sales to potential adopters. Third, only a limited number of studies have integrated predictive accuracy benchmarking and country-level trajectory clustering into a single framework. Most forecasting studies have emphasized performance metrics for an algorithm while diffusion studies have emphasized factors influencing an individual's decision to adopt; therefore, there has yet to be sufficient examination of the synergy between the two streams of research [3]. Fourth, researchers have thus far not adequately assessed the extent to which utilizing only historical sales patterns and regional encoding to forecast EV sales will accurately capture how well algorithms will predict long-term performance by using only signals from the sales data's own temporal and spatial characteristics. All these combined issues suggest the need for a more comprehensive comparative study utilizing seven regression algorithms across multiple countries and performing integrated clustering analysis on all data sets. Therefore, because of this, the predictive capabilities of conventional forecasting methodologies have been lacking.

2 Methodology

In order to understand how the forecasted results were achieved, the theoretical formulations that form the backbone of the algorithms will be explained in terms of their mathematical foundations.

The International Energy Agency (IEA) compiled the dataset through contributions and reports submitted by member countries and industry bodies, as well as from its own analytical models. Therefore, every piece of data has gone through extensive cross-verification, and updates occur once per year. The dataset provided a comprehensive overview of the electric vehicle transition, including historical numbers for EV sales, stock, and market share, as well as long-term forecasts for all types of light-duty vehicles, including battery electric, as well as medium- and heavy-duty commercial vehicles developed using plug-in technology, across 40+ countries and different geographic regions. This information can be used to understand how quickly electric vehicles are being adopted; assess how much countries are adopting electric vehicles; assess how ambitious their future targets are; and build their own forecast models on top of one of the most reliable data sets in existence. If you are using this information to publish findings, you should always cite the source of your data as the IEA Global EV Outlook 2025, and where possible, set up a link to the IEA's Global EV Data Explorer. The electric vehicle (EV) sales dataset contains historical data on EV sales across multiple countries/regions between 2010 and 2024 and will include time variables, geographical indicators, and quantitative indicators. The three types of variables above will be used as a basis for descriptive and predictive analysis.

After filtering, the analysis file had a total of 648 records (rows), covering 43 countries/regions over 15 years, from 2010–2024. Although the original data provides yearly totals at the country-level, they cannot provide any monthly or quarterly disaggregation, therefore each observation is yearly. The resulting table was converted to a wide format suitable for supervised learning, where each country had 2010–2020 sales values as training set, and the 2021-2024 sales value was the test set. This means that there were 43 samples (countries) and 15 predictor features; 14 years of historical sales plus region identifier that was label-encoded. To correct the exponential growth curve of sales, a natural-log transformation was applied to all sales values to stabilize the variance. These preprocessing methods ultimately provided a concise numeric feature table ready for model training; however, though there were only 43 countries to train on, there were still only 15 features available for training.

Temporal trends were evaluated through time series plots, but the data will show a generally exponential rise in EV adoption within most of the regions where data were recorded. Comparative analysis was conducted between individual countries to determine which countries are leading and lagging in terms of adoption rates.

The CV measures of growth, including year-over-year and compound annual growth rates, will also be calculated beginning with the following expression [11]:

$$CAGR = \left(\frac{V_f}{V_i}\right)^{\frac{1}{n}} - 1 \quad (1)$$

Where:

V_f is the final value of the item or assets being measured,

V_i is the initial value of the item or assets being measured,

n is the number of years for the investment period.

This metric gives a smoothed out annualized growth rate for measuring the long-term growth trends associated with an investment.

2.1 Feature engineering

One of the most important ways to improve the prediction performance of the models was using feature engineering. The historical sales data for selected years was used as the input variables to predict the sales pattern of EVs from 2021 to 2024. Since the growth of EV adoption occurs exponentially, we applied a Log Transformation to stabilize the variance and create linear relationships between the input and predicted variables.

This transformation is illustrated below:

$$x' = ln(1 + x) \quad (2)$$

The effect of outliers on the data has been reduced because of this transformation and the resulting normality in distribution has allowed for improved model performance [12]. Categorical variables have also been transformed to numerical representations via label encoding. This process assigns an integer to each category so that machine learning models can quantify the differences between regions.

2.2 Model development

The framework for developing predictive models for EV sales was created as a supervised regression type, which is defined as predicting future sales of electric vehicles using historical data. To accomplish this, we have deployed multiple regression techniques and machine learning techniques to minimize biases in model results and ensure robustness.

It utilizes an end-to-end system comprised of four distinct phases: data acquisition and preparation; log transformation and label encoding; model development (through the use of regression analysis) and evaluation; and model selection via a "best performer" decision node. Thereafter, the workflow splits, with one branch used for forecasting 2025-2028 electric vehicle (EV) sales while the other branch performing dimensionality reduction and clustering for the identification of country clusters through principal component analysis (PCA) and trajectory-based grouping. This results in the provision of both predictive sales forecasts and policy-relevant country classification related to conclusions drawn from this analysis.

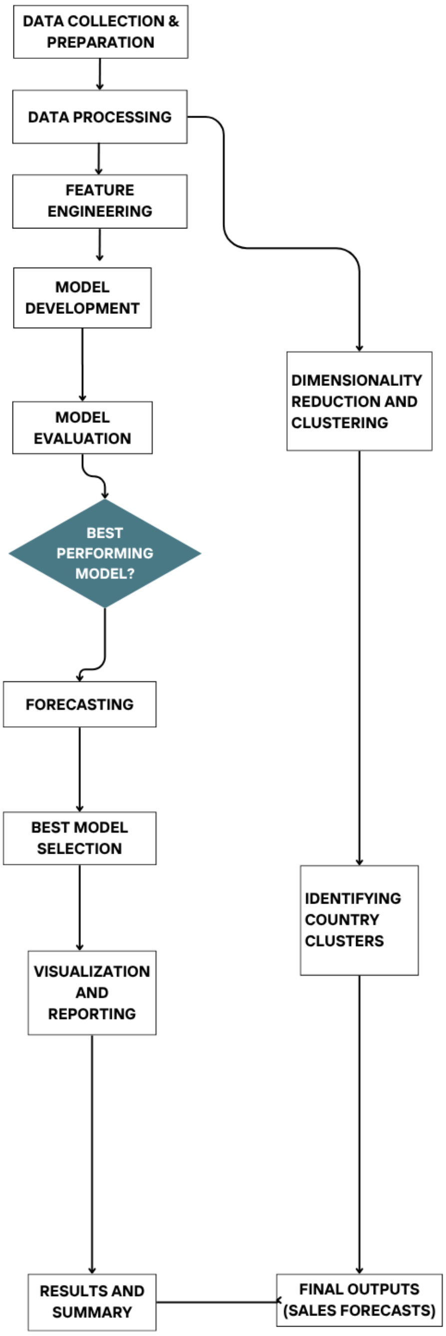

Fig. 1. Analytical pipeline

2.3 PCA Component fomulation

Dimensionality reduction and PCA were applied to all the time series' data from 2010–2024 for the 43 Countries. Each country was represented as a matrix of its EV sales for the 15 years between 2010 and 2024 (43 × 15 matrix). Since the sales data for adjacent years is highly correlated, the 15-year data were reduced to a 2-dimensional space using PCA where the majority of variation in sales can still be represented. The first two principal components (PCs) together account for over 90% of the total variance in the data from the 15 years. The first PC measures each country's total sales volume and cumulative sales growth. The second PC separates countries based on the time at which they began accelerating their adoption; this allows for a distinction to be made between early adopters and late adopters. PCA thus made it possible to perform k-means clustering on data that had been reduced in noise and was structured in meaningful ways resulting in the four groupings of countries shown in the final country-cluster plot.

Given the centered data matrix X∈R^(n×d) with n countries and d annual time points, PCA seeks a set of orthogonal directions w_k that maximize the variance of the projected data [13]:

$$w_k = \underset{\|w\|=1}{\arg\max} Var(Xw) \;\; subject\ to\ w_k \perp w_1, \dots, w_{k-1} \tag{3}$$

These directions are the eigenvectors of the covariance matrix

$$C = \frac{1}{n-1} X^T X, \tag{4}$$

ordered by devaluing eigenvalues λ_k. The proportion of variance by the first k components would then be

$$\sum_{i=1}^{k} \lambda_i / \sum_{i=1}^{d} \lambda_i. \tag{5}$$

For clustering, each country's 15-year trajectory was projected onto the first two principal components via

$$Z = XW_2 \tag{6}$$

Where W_2 is the $d \times 2$ matrix of the first two eigenvectors, yielding a two-dimensional representation used for k-means.

2.4 Model evaluation

Model performance was evaluated using common regression metrics. The MAE (Mean Absolute Error) counts average size of errors, while the RMSE (Root Mean Square Error) puts more weight on bigger errors. The R^2 was the main metric used to determine the best performing model R^2 defined by the following:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \tag{7}$$

MAE and RMSE are gotten through the following equations [13]:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |e_i| \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} e_i^2} \tag{9}$$

Therefore, the model which has the highest R^2 and smallest error metrics will be used.

3 Results

3.1 Global EV trends

The global electric vehicle market has been evolving at a remarkable rate between 2010-2024 with notable trends emerging. In 2010, the global sales of electric vehicles were only 0.01 million units; however, they grew steadily over the first few years to reach 0.20 million by 2014. After 2015, there was an observable acceleration in the growth of new electric vehicle sales, with sales increasing to 0.47 million units in 2016 and surpassing 1 million units by 2018. The growth has continued and accelerated dramatically from 2020 onward, with sales increasing from 2.00 million in 2020 to 4.70 million in 2021, 7.30 million in 2022, 9.50 million in 2023, and 11.00 million in 2024. As a result, the general shape of the curve is that of an exponential function, with the last four years showing the largest increase in this trend, clearly highlighting the rapid global transition to electric mobility.

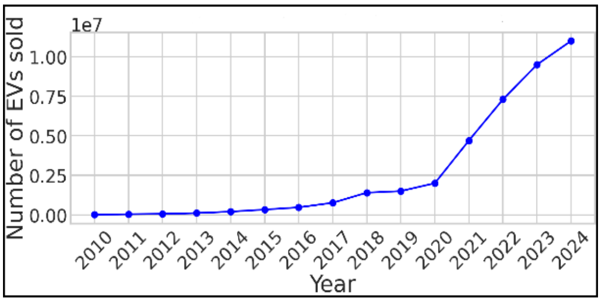

Fig. 2. Global EV sales (2010-2024

Looking at the historical breakdown of the market share over the period between 2010 and 2024 by year for the five largest markets, in the first few years, 2010 to 2013, China's share fluctuated between 20% and 35%, while the USA was between 10% and 20%. However, starting in 2014, China's share increased dramatically, reaching a peak of 200% in 2021 and subsequently stabilizing around 40% by 2024. Conversely, the USA's share decreased from around 50% in 2015 and 2016 to approximately 40% in 2024. Germany has had a modest market share that has fluctuated between 5% and 20%, while the UK and France have also averaged around 5% to 15%.

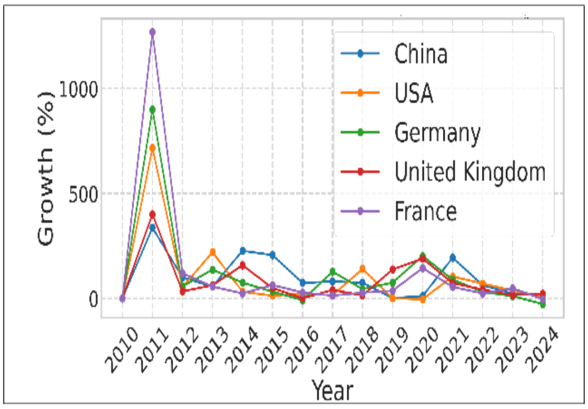

Fig. 3. Year-over-Year EV sales growth (%) for Top 5 Countries

3.2 The regression models

A regression analysis of seven different models was conducted to predict 2024 EV sales based upon historical sales data and the region in which those sales occurred. The following models were investigated: linear regression, ridge regression, random forest, gradient boosting, XGBoost, k-nearest neighbors (KNN), and support vector regression (SVR). Each model was evaluated using three performance metrics, namely: mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination R^2. The investigation yielded very different estimates of electric vehicle sales in 2024 between algorithms. Only linear and ridge regression had extremely low values of MAE, RMSE, and R^2. R^2 values less than zero indicate that the models are worse at predicting sales. Therefore, linear based approaches to predict 2024 sales do not give an accurate estimate of the nonlinear relationship between the features and 2024 EV sales.

Table 1. Regression models results

Model	MAE	RMSE	R^2
Ridge	2038.911	2515.457	-1.494
Linear Regression	2243.447	2823.297	-1.503
XGBOOST	112827.106	122356.754	-4.663
Gradient Boosting	113153.631	122422.418	-4.867
Random Forest	116833.631	126078.502	-5.849
KNN	127195.527	135910.979	-8.267
SVR	144095.821	152007.314	-11.039

The tree-based and distance-based models (XGBoost, Gradient Boosting, Random Forest, K-Nearest Neighbours (KNN), Support Vector Regression (SVR)) produced higher MAE values (most exceeding 100,000) and RMSE values (most exceeding 120,000) than the linear regression models. The higher error values indicate that these models overfitted the training data, as well as could not adjust to the limited sample data and the exponential growth of EV adoption. The negative R^2 values for these models (-4.6 to -11.0) indicate how much these models' forecasts differed from the actual sales of EVs from 2021 to 2024, particularly in countries with either rapid EV sales growth or sudden EV market saturation.

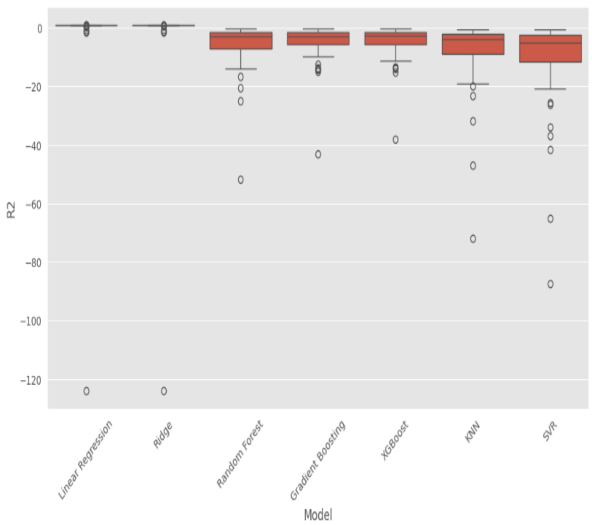

Fig. 4. Distribution of R^2 scores accros countries

Overall, ridge regression was the best of the algorithms considered, producing the lowest mean absolute error, root mean square error and least negative R2 of all the evaluated methods. Although ridge regression had the lowest predictive accuracy of the models tested, the regularization employed by ridge regression was successful at controlling variance and producing the most consistent results across countries. Therefore, the results indicate that simpler linear models with regularization will outperform complex ensemble or kernel-based techniques for short-term forecasting of EV sales trajectories given the volatility and size of the data set.

3.3 Predictions

3.3.1 EV sales (comparison of regression models behavior)

Using the actual EV sales data from 2021-2024 for the top 5 countries, we compared their sales data to the forecasts generated by each of the 7 regression models. With the predicted trajectories being plotted with each country's observed sales, we were able to visually and statistically evaluate how closely the predicted values matched actual sales. Each model demonstrated different patterns or levels of correspondence between the forecasts and the actual EV sales by country. The Linear and Ridge Regression models in China, the United States, and the United Kingdom matched the increase in actual EV sales over time and clearly show how steady growth occurred between 2021-2024. In fact, they showed a general upward trend for all three countries relative to observed data, which indicates that simple linear functions or equations between time features and EV sales had a high level of precision for predicting EV adoption on a short-term basis. By contrast, the Random Forest, Gradient Boosting and XGBoost ensemble forecasting models showed similar deficiencies in their predictions; instead, they tended to produce flattening forecast trends. This indicates that these models had difficulty extrapolating beyond available historical data of EV sales based on slow-to-modest growth trends experienced prior to 2020, because such slow-to-moderate growth rates had been common in training data collected from 2010 through 2020.

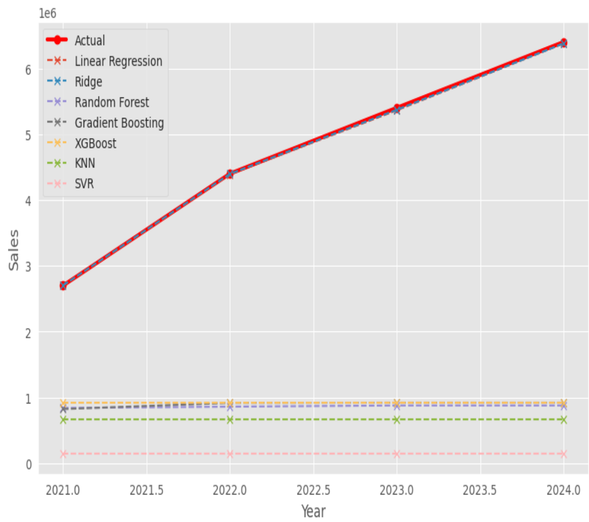

Fig. 5. EV forecast sales vs actual (2021-2024) - China

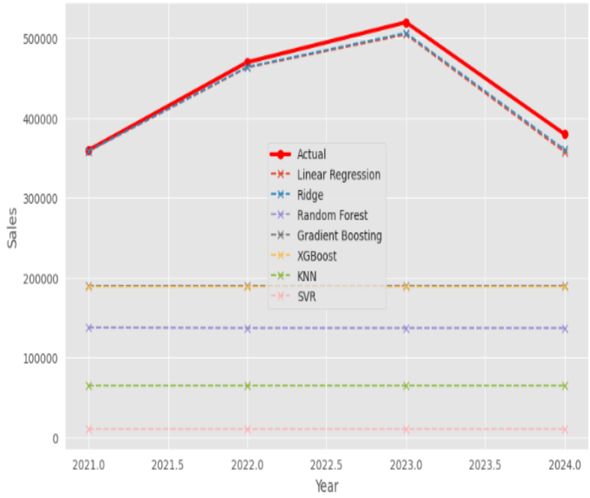

Fig. 6. EV forecast sales vs actual (2021-2024) - Germany

The pattern of smooth curves produced by Ridge and Linear Regression matched actual sales patterns in Europe (Germany and France) while the complex models essentially produced flat lines on the projections. The KNN and SVR models performed the worst. They produced near horizontal line on the forecast horizon, as their algorithms require a heavy reliance on local data patterns and distance metrics, which are poor fits for time series with exponential growth. The KNN and SVR models were unable to capture the long-term momentum of demand nor the long-term dependency that are critical to analytics resulting in a significant underestimate of sales volume. Ensemble Models, the most effective in theory, were hampered by the limitations of the sample leverage and the omission of external explanatory variables influencing electric vehicle faster or establish electric vehicle faster.

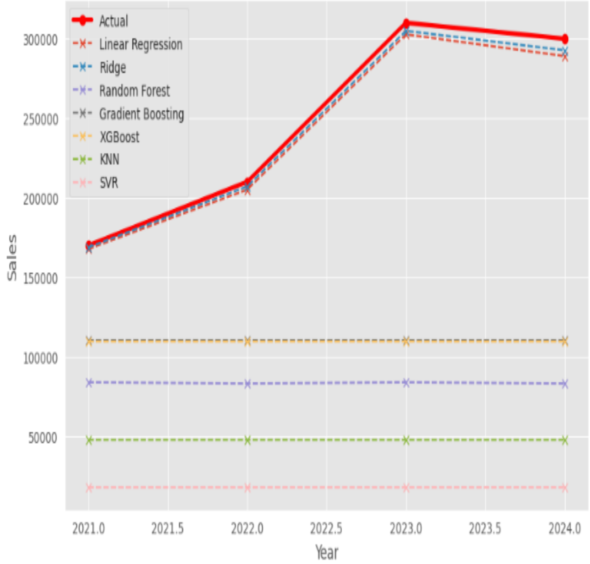

Fig. 7. EV sales forecast for top 5 countries

Ridge regression was an effective solution that modelled the data accurately by addressing both bias and variance through the regularization applied, which allowed for classification and prediction, through the use of regularized coefficient values. In return for being penalized for large values of coefficients, Ridge effectively reduced the effect of collinearity of variables on lagged sales and rolling averages as well as minimized the effects of random noise or extreme year-on-year variability; thereby creating a positive relationship between bias and variance and enabling generalization across countries with different speeds of adoption. Overall, Ridge was able to model the principal form of the global EV growth trend without being affected by short-term trends in the future, so being the best algorithm to predict beyond the algorithms tested.

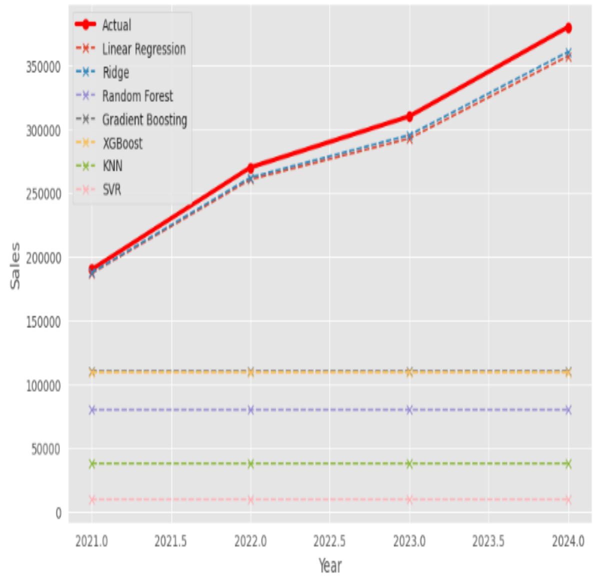

Fig. 8. EV forecast sales vs actual (2021-2024) - UK

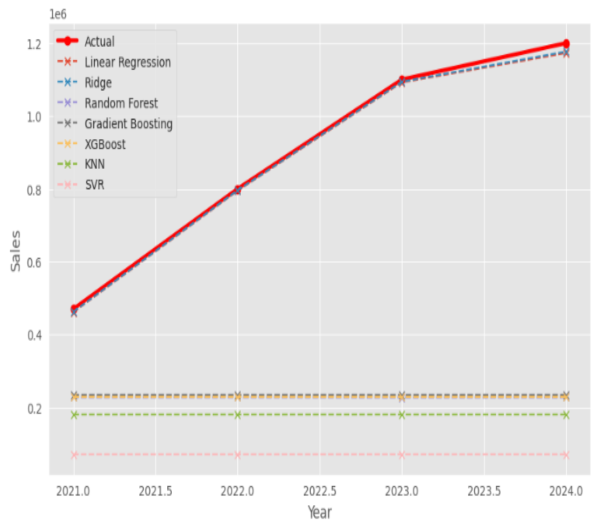

Fig. 9. EV forecast sales vs actual (2021-2024) - USA

3.3.2 EV sales Trajectory

The leading two axes of clusters, PC1 and PC2, explain the primary variation in historical EV sales growth among different nations worldwide. There are four identifiable cluster groups, colour coded green (cluster 0), orange (cluster 1), blue (cluster 2) and purple (cluster 3), where the points plotted along the two axes in the chart depict countries will differ based upon both the scale and the shape of the country's EV adoption curve. Countries like South Africa, Colombia and Estonia make up cluster 0.

Most of the countries within cluster 0 are emerging or middle income and have relatively low historical EV sales, but have had rapid increases in their sales in the last few years and therefore they are viewed as potentially high-growth markets during the forecast period. Countries like China, the UK, France, Canada, Sweden, the Netherlands and Norway make up the countries in cluster 1. Cluster 1 is comprised of countries that have mature and high-volume EV sales but that are showing early adoption, constant growth and in some instances, saturation (e.g. Norway). Cluster 1 countries have trajectories with positive trending, but also suggest an annual decrease in the growth of the market as well, therefore indicating the national market is reaching their peak in sales made. Indonesia, Vietnam, Seychelles, form cluster 3, representing both advanced economies in Asia and countries in Southeast Asia that are experiencing rapid economic development.

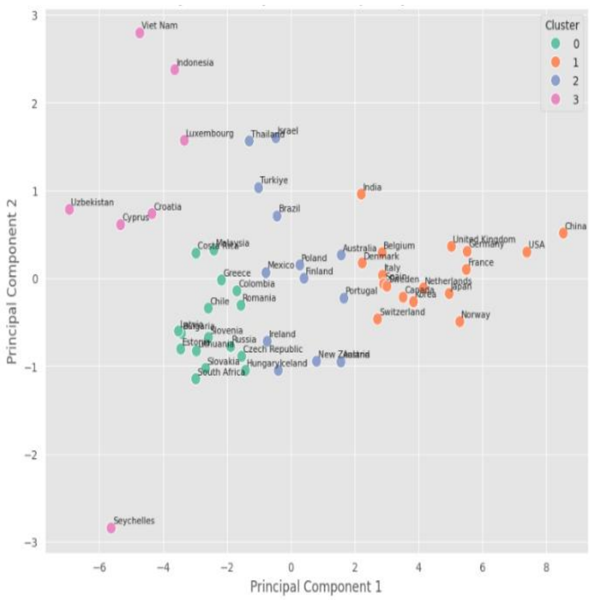

Fig. 10. Country clusters by EV sales trajectory

There is considerable heterogeneity across the two groups regarding their growth paths, with certain countries experiencing consistent moderate growth rates, while others have only recently accelerated their growth rates from a lower baseline. The location of the points on PC2 indicates that there may be significant differences across countries in terms of when they began to undertake the first steps in their transformation. The PCA plot effectively reduces the dimensionality of the time series data, showing that PC1 primarily separates countries based on total sales and total growth and that PC2 primarily separates early adopting and late starting countries. Clustering demonstrates that there is a consistent core-periphery relationship throughout the world regarding EV adoption, and there is a distinct line drawn between established leaders, rapidly developing economies, and very low volume outliers

4 Conclusion

While the overall trend of global EV sales follows an exponentially growing pattern, the majority of models we tested were not able to appropriately represent this type of growth, as tree-based ensemble and distance-based models generated extremely high error and negative R^2 results, indicating an inability to generalize the COVID-19-related highs and lows of slow growth from the training dataset. Linear models, while missing some of the more complex dynamics of the sales trend, produced much more stable forecasts and avoided producing any extreme outliers compared to the more complex models.

Among all seven different models, ridge regression constantly generated the best results—lowest mean absolute and root mean square errors, and least negative R^2 values. Ridge regression's use of regularization provided a good balance between bias and variance, allowing it to mitigate the impact of collinearity related to lagged features and smooth out the influence of random noise present in the datasets. As a result, ridge regression is more resilient to short-term market changes and better suited to short-term forecasting than any of the other models we tested, even though it does not fully capture the exponential growth trend associated with EV adoption.

In summary, the findings indicate that on smaller datasets with high volatility (for example, the country-level EV sales from 2010 to 2024), simpler and regularized techniques such as linear regression may provide better performance than complex techniques such as ensembles or kernels. However, for more extended forecasts (e.g., out to 2030), a combined approach using Ridge regression along with other types of non-linear regression or time-series techniques can give a more accurate representation of the adoption curves in the market. By using both methods, they will be able to take advantage of Ridge's stability while also considering the exponential growth trends that characterize the global electric vehicle transition.

Data Availability: For this analysis, the Global Electric Vehicle Dataset 2025 was used. It was obtained via the Global EV Data Explorer from the International Energy Agency (IEA) https://www.kaggle.com/datasets/padmapiyush/global-electric-vehicle-dataset-2023?resource=download.he

References

1. D. Li, A.D. Lau, Y. Gong, Electric vehicles empowering the construction of green sustainable transportation networks in Chinese cities, Dynamic evolution frontier trends and construction pathways, Energies 18, 8 (2025) 19-43.
2. B. Srinivasan, Predicting future electric vehicle sales, A time series forecasting approach using historical EV sales data, International journal for applied information management 5, 3 (2025) 177-189.

3. G. Han, et. al., A structured review of electric vehicle sales research, Multi-level driving factors and forecasting pathways over the past decade, World electric vehicle journal 17, 3, (2026) 122.
4. A. Ajiono, T. Hariguna, Comparison of three time series forecasting methods on linear regression, Exponential smoothing and weighted moving average, International Journal of Informatics and Information Systems 6, 2 (2023) 89–102.
5. S.K. Shil, M.S. Chowdhury, N.R. Tannier, M.T. Tarafder, R. Akter, N. Gurung, M.M. Sizan, Forecasting electric vehicle adoption in the USA using machine learning models, Journal of computer science and technology studies 6, 5 (2024) 61–74.
6. S. Dirir, Predictive modeling of electric vehicle adoption in OECD nations, An empirical assessment, Quality & quantity (2026) DOI: 10.1007/s11135-y
7. S. Veysi, M. Moshfeghi, A. Sadrfaridpour, P. Emamy, Electric vehicle sales forecast for the UK, Integrating machine learning time Series Models and global trends, Algorithms 18, 7 (2025) 430 DOI: 10.3390/a18070430.
8. S. Afandizadeh, D. Sharifi, N. Kalantari, H. Mirzahossein, Using machine learning methods to predict electric vehicles penetration in the automotive market, Sci. rep. (2023) doi: 10.1038/s41598-023-35366-3.
9. C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, A comparative analysis of gradient boosting algorithms, Artificial Intelligence Review 54 (2021) 1937–1967.
10. T. Chen, C. Guestrin, XGBoost, A scalable tree boosting system, Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining San Francisco CA ACM (2016) 785–794.
11. L. Maybury, P. Corcorana, L. Cipcigan, Mathematical modelling of electric vehicle adoption, A systematic literature review, Transport and environment 107 (2022) 103278.
12. M.P. Clements, P.H. Franses, and N.R. Swanson, Forecasting economic and financial time series with non-linear models, International journal of forecasting 20, 2 (2004) 169–183.
13. F.L. Gewers et al., Principal component analysis, A natural approach to data exploration, ACM Computing Surveys CSUR (2021) 1-34.