Automobile Valuation through Supervised LearningTechniques:AComparativeAnalysis

Anuradha M Sandi1*,AmulyaRatna[2], Swati[3]

DepartmentofComputerScience and EngineeringGuruNanakDevEngineeringCollege Bidar, India

ratnaamulya.j154@gmail.com

Abstract—

The pre-owned automobile sector has witnessed rapid expansion globally, necessitating accurate and reliable pricing mechanisms. This research investigates the application of diverse supervised learning algorithms to forecast automobile valuations based on multifaceted vehicular attributes. Seven distinct regression-based predictive models were constructed and systematically evaluated: Multiple Linear Regression (MLR), Lasso Regression, Ridge Regression, Elastic Net Regression, Decision Tree Regressor, RandomForestRegressor,andSupportVectorRegressor(SVR).To mitigate overfitting and improve generalization capability, a structured approach incorporating regularization strategies and systematic hyperparameter optimization via GridSearchCV with 10-fold cross-validation was adopted. The predictive models were assessedonapubliclyaccessibleautomotivedatasetencompassing attributes such as vehicle age, fuel economy, engine displacement, road tax, and fuel category. Empirical findings demonstrate thatthe Support Vector Regressor, configured with an RBF kernel and optimized hyperparameters, achieved superior predictive performance with an R^2 value of 95.83%, Mean Absolute Error (MAE) of 0.136, Mean Squared Error (MSE) of 0.042, and Root Mean Squared Error (RMSE) of 0.204 on the 90:10 train- validation partition. The outcomes of this investigation provide actionable insights for buyers, sellers, and automotive platforms seeking data-driven valuation tools.

Keywords— *Automobile valuation, Supervised learning, Support Vector Regressor, Hyperparameter optimization, Regression analysis, Predictive modeling, Cross-validation.*

I. INTRODUCTION

The global pre-owned automobile industry has undergone a significant transformation over the past decade, evolvinginto one of the most rapidly advancing commercial sectors worldwide. Market analysis reports indicate that the valuation of this industry has experienced substantialgrowth, driven by increased consumer preference for cost- effective vehicle acquisition options. The proliferation of digital automotive marketplaces has further accelerated this trend, providing both purchasers and vendors with unprecedented access to comprehensive vehicle information and pricing trends [1].

Determining an equitable market valuation for a pre-owned vehicle remains a complex undertaking, influencedby numerous interdependent parameters including the vehicle's chronological age, cumulative distance traveled, engine specifications, fuel economy metrics, and prevailing market conditions. Conventional valuation methodologies often rely on subjective assessments and rudimentary heuristic approaches, which frequently produce inconsistent and unreliable estimates [2]. This inconsistency creates significant challenges for both parties in a transaction—purchasers risk overpaying while sellers may undervalue their assets.

Machine learning (ML) algorithms present a robust computational framework for addressing this valuation challenge by identifying intricate non-linear relationships withinmultivariateautomotivedatasets.Supervisedlearning paradigms, specifically regression-based approaches, have demonstrated considerable promise in constructing accurate predictive models that can estimate vehicle prices with high fidelity [3]. By leveraging historical transactional data alongside technical vehicle specifications, these algorithms can capture complex feature interactions that traditional statistical methods may overlook.

This research undertakes a systematic comparative evaluation of seven supervised regression algorithms to identify the most effective approach for automobile price forecasting. The investigation addresses two primary challenges: (i) preventing model overfitting through the application of regularization techniques, and (ii) optimizing model parameters using structured hyperparameter search methodologies. The study contributes a comprehensive benchmarking framework that evaluates each algorithm across multiple performance dimensions using standardized evaluation metrics.

The remainder of this manuscript is structured asfollows: Section II presents a critical review of related literature. Section III details the methodological framework including data preprocessing, algorithm descriptions, and evaluation criteria. Section IV outlines the experimental configuration and hyperparameter optimization procedures. SectionVdiscussestheempiricalfindingsandcomparative analysis. Section VI concludes with key insights and future research directions.

II. LITERATUREREVIEW

Numerous investigations have explored the utilization of computational intelligence techniques for vehicle price estimation. This section presents a critical synthesis of significant prior works, highlighting their methodological approaches, key findings, and identified limitations.

Gegicetal.[4]constructedpredictionmodelsemploying Support Vector Machines (SVM) and Artificial Neural Networks (ANN) to estimate automobile valuations. Their SVM implementation achieved an R^2 coefficient of 90%, while the ANN model attained 85.71%. Although theauthors provided a thorough analysis of their experimental outcomes, the study lacked a systematic comparison with contemporary approaches, limiting the contextualization of their findings within the broader research landscape.

Asghar et al. [5] applied the Ordinary Least Squares (OLS) regression framework augmented with the Variance Inflation Factor (VIF) methodology to develop forecasting models. Their Lasso regression variant achieved an R^2metric of 91%. While the methodological pipeline was articulated with exceptional clarity, the investigation did not present disaggregated accuracy measurements for individual algorithmic variants, making it challenging to assess the relative contribution of each modeling choice.

Venkatasubbu and Ganesh [6] utilized supervised learning paradigms to construct automobile retail price estimators. Their experimental design incorporated Lasso Regression, Multiple Linear Regression, and Decision Tree algorithms, achieving error rates of 3.58, 3.46, and 3.51 respectively. The authors employed rigorous statistical validation procedures including Analysis of Variance (ANOVA); however, the exclusive reporting of error rates without complementary goodness-of-fit statistics limited the comprehensive assessment of model adequacy.

Gajera et al. [8] conducted an extensive comparative study incorporating Random Forest, K-Nearest Neighbors (KNN), Decision Tree, XGBoost, and Linear Regression algorithms. The Random Forest implementation demonstrated the highest predictive capability with an R^2 value of 93.11%. Their analysis was thorough in its algorithmic coverage; however, the absence of detailed evaluation metrics beyond the coefficient of determination represents a notable gap in their reporting.

A synthesis of the reviewed literature reveals that while substantial progress has been made, several research gaps persist. These include limited exploration of systematic hyperparameter tuning, insufficient comparative analysis across diverse train-test partitioning strategies, and the need for more comprehensive evaluation using multiple complementary error metrics. This study addresses these identified gaps through a rigorous multi-model comparison with optimized hyperparameters and multi-dimensional performance assessment.

TABLEI:COMPARATIVESUMMARYOFPRIORRESEARCHONAUTOMOBILEPRICEPREDICTION

Ref.	Methodology	ResearchObjective	Algorithms Employed	R^2(%)	Strengths	Limitations
[4]	Regression-based MLapproaches	Optimal prediction modelidentification	SVM,ANN	90.00,85.71	Comprehensiveresult analysis	Nocomparisonwith existing research
[5]	OLSwithVIF selection	Pre-ownedvehicle price estimation	Lasso,Linear Regression	91.00	Clearmethodological exposition	Individual model metricsunreported
[6]	SupervisedML techniques	Retailpricingmodel development	Lasso,MLR, DecisionTree	ErrorRate:3.46–3.58	Statisticalvalidation (ANOVA)	Onlyerrorrates reported
[7]	SupervisedMLwith regularization	Priceestimationmodel development	Linear,Ridge, Lasso	Notspecified	Well-documented flowchart	Noquantitative results
[8]	Linearandensemble models	ML-basedprice predictor	RF,KNN,DT, XGBoost,LR	93.11,69.66, 84.30,92.04, 76.46	Thoroughanalysisand comparison	Incomplete evaluationmetrics

III. METHODOLOGY

The proposed methodological framework comprises five sequentialstages:datasetacquisition,datapreprocessingand transformation, model construction, model assessment, and target variable prediction. Figure 1 presents the architectural overview of the complete analytical pipeline.

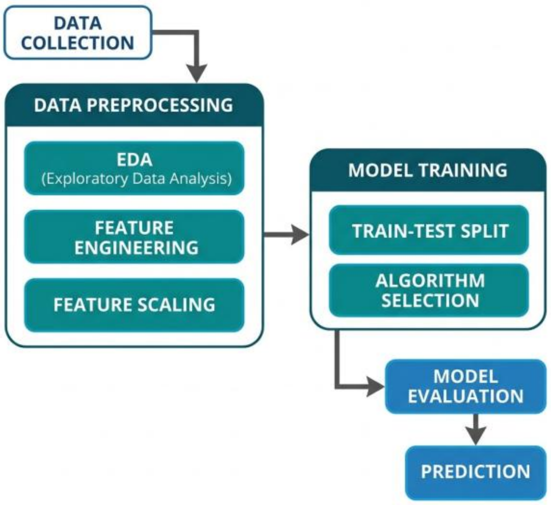

Fig.1. Architecturaloverviewoftheproposedmachinelearning pipeline for automobile valuation

A. DatasetAcquisition

The experimental dataset was sourced from the Kaggle repository [20], comprising comprehensive records of pre-owned vehicles with diverse technical and commercial attributes. The dataset encompasses vehicular features including manufacturing year, selling price, kilometers driven, fuel category, transmission type, enginedisplacement (in CC), maximum power output (in BHP),andapplicableroadtax.Thedatasetcontained8,128records with 13 attribute columns, providing sufficient volume for robust model training and validation.

TABLEII: DATASETATTRIBUTE SUMMARY

Sr. No.	Feature Name	Data Type	Description
1	Year	Integer	Manufacturingyearof vehicle
2	Price	Float	Sellingprice(target variable)
3	Mileage	Float	Fueleconomy(km/l)
4	EngineSize	Integer	EnginedisplacementinCC
5	MaxPower	Float	MaximumpowerinBHP
6	FuelType	Categorical	Petrol/Diesel/CNG/ Electric
7	Transmission	Categorical	Manual/Automatic
8	RoadTax	Float	Applicableroadtaxamount

B. DataPreprocessingandTransformation

Data preprocessing constitutes a critical preparatory phase that directly influences the accuracy and reliability of ML models. The preprocessing pipeline, illustrated in Figure 2, consists of three principal operations: categorical variable encoding, data cleansing, and feature normalization.

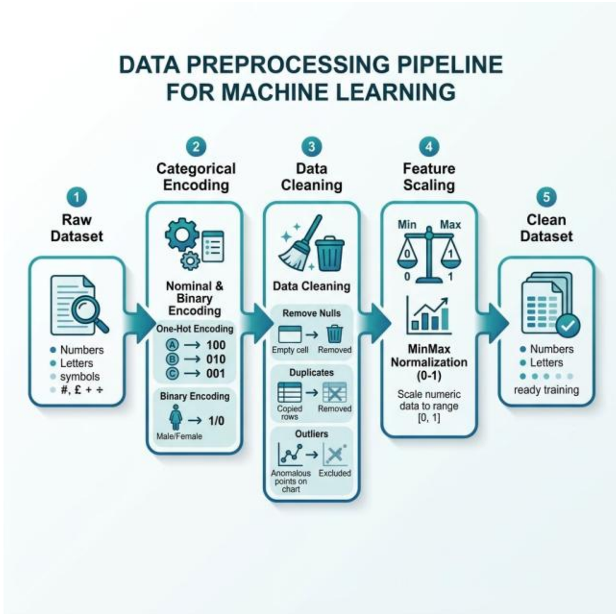

Fig.2. Datapreprocessingpipelineillustratingthesequential transformation stages

Categorical Variable Encoding: Non-numerical attributes were transformed through a two-stage encoding strategy. Initially, nominal encoding [9] was applied to assign unique integer identifiers to categorical values. Subsequently, binary encoding [10] was employed to represent these integers in binary format, distributing the encoded values across multiple columns. This approach significantly reduces dimensionality compared to one-hot encoding while maintaining the representational integrity of categorical features.

DataCleansing: Thedatasetunderwentrigorousquality assessment to identify and rectify inconsistencies. Records containing null values, duplicate entries, and corrupt data points were systematically identified and removed. Outlier detection was performed using the Interquartile Range(IQR) method, enabling the removal of extreme values that could disproportionately influence model training.

Feature Normalization: To ensure commensurability across features with disparate scales, the Min-Max normalization technique [11] was applied, constraining all feature values within the [0, 1] interval. This transformation is mathematically expressed as:

$$X_{normalized}=(X-X_{min})/(X_{max}-X_{min}) \tag{1}$$

where X represents the original feature value, and $X_{min}$and X_{max}denote the minimum and maximum values of that featureacrossthedatasetrespectively.Thisnormalization mitigates the dominance of high-magnitude features during model optimization.

C. ModelConstruction

The preprocessed dataset was partitioned into training and validation subsets at ratios of 90:10, 80:20, and 70:30 respectively.Sevenregressionalgorithmswereimplemented andtrainedonthedesignatedtrainingpartitions:

1) MultipleLinearRegression(MLR): MLRmodelsthe linear association between multiple independent predictors ($x_1, x_2, ..., x_n$) and a continuous dependent variable (y). The modelestimatestheparameters$\beta_0,\beta_1,...,\beta_n$thatminimize thesumofsquaredresiduals[12]:

$$y=\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_nx_n+\varepsilon \quad (2)$$

2) RidgeRegression(L2Regularization): Ridge regressionextendsordinaryleastsquaresbyincorporatinga penalty term proportional to the sum of squared coefficient magnitudes.This L2 penalty constrains the parameter space, mitigatingmulticollinearityandreducingmodelvariance [13]:

$$L_{Ridge}=\Sigma(y_i-\hat{y}_i)^2+\lambda\Sigma(\beta_j^2) \quad (3)$$

where λ is the regularization parameter controlling the penalty intensity.

3) Lasso Regression (L1 Regularization): The Least Absolute Shrinkage and Selection Operator applies an L1 penalty to the regression coefficients, which encourages sparsity by driving irrelevant coefficient estimates toward zero [14]:

$$L_{Lasso}=\Sigma(y_i-\hat{y}_i)^2+\lambda\Sigma|\beta_j| \quad (4)$$

4) ElasticNetRegression: ElasticNetcombinesbothL1 and L2 regularization terms, offering a balanced approach that retains the feature selection capability of Lasso while maintaining Ridge's handling of correlated predictors [15]:

$$L_{ElasticNet}=\Sigma(y_i-\hat{y}_i)^2+\lambda_1\Sigma|\beta_j|+\lambda_2\Sigma(\beta_j^2) \quad (5)$$

5) Support Vector Regressor (SVR): SVR extends the margin-based classification principles of Support Vector Machines to regression tasks. The algorithm identifies an optimal hyperplane that maximizes the number of training instances within an ε-insensitive tube around the prediction boundary. The radial basis function (RBF) kernel enables SVR to capture non-linear relationships in the feature space [16].

6) Random Forest Regressor: This ensemble methodology constructs multiple independent decision trees on bootstrapped subsets of the training data and inherent randomization in both sample and feature selection reduces variance and suppresses overfitting tendencies [17].

7) Decision Tree Regressor: Decision trees recursively partitionthefeaturespacebasedonoptimalsplittingcriteria, creating hierarchical decision structures where each leaf node represents a predicted value. The Friedman MSE criterion was employed for split quality assessment [18].

D. ModelEvaluationMetrics

Fourcomplementaryperformanceindicatorswereemployed to quantify predictive accuracy [19]:

1) Coefficient of Determination (R^2): quantifies the proportion of variance in the dependent variable that is predictable from the independent variables:

$$R^2=1-[\Sigma(y_i-\hat{y}_i)^2/\Sigma(y_i-\bar{y})^2] \quad (6)$$

2) Mean Absolute Error (MAE): computes the average magnitude of prediction deviations without considering directionality:

$$MAE=(1/n)\Sigma|y_i-\hat{y}_i| \quad (7)$$

3) MeanSquaredError(MSE): amplifieslarger prediction errors through quadratic penalization:

$$MSE=(1/n)\Sigma(y_i-\hat{y}_i)^2 \quad (8)$$

4)RootMeanSquaredError(RMSE): providesan interpretable error measure in the original unit scale:

$$RMSE=\sqrt{[(1/n)\Sigma(y_i-\hat{y}_i)^2]} \quad (9)$$

IV. EXPERIMENTAL CONFIGURATION AND HYPERPARAMETEROPTIMIZATION

All experiments were conducted within the Jupyter Notebook integrated development environment executing Python 3.10.6. The computational infrastructure utilized for this investigation is summarized in Table III.

TABLEIII: COMPUTATIONALINFRASTRUCTURE SPECIFICATIONS

Component	Specification
Processor	IntelCorei5-1135G7@2.40GHz(11th Gen)
SystemMemory(RAM)	16GBDDR4
GraphicsProcessor	IntelIrisXe,2GB
OperatingSystem	MicrosoftWindows11(64-bit)
Development Environment	JupyterNotebook
ProgrammingLanguage	Python3.10.6
MLLibraries	Scikit-learn,NumPy,Pandas

Hyperparameter optimization was performed using the GridSearchCV exhaustive search methodology with 10-fold cross-validation to identify the parameter configuration yielding maximum predictive performance. The search process systematically explores a predefined parameter grid, evaluatingeachcombinationthroughcross-validatedscoring topreventselectionbias.TableIVenumeratestheoptimized hyperparameters obtained for each algorithm

TABLE IV: OPTIMIZED HYPERPARAMETER CONFIGURATIONSFORREGRESSIONMODELS

V.

Sr. No.	Algorithm	Hyperparameter	OptimalValue
1	RidgeRegression	alpha(λ)	0.01
2	LassoRegression	alpha(λ)	0.001
3	ElasticNet	alpha(λ)	0.0001
4	SVR	degree	1
		gamma	auto
		kernel	rbf
5	DecisionTree	criterion	friedman_mse
		splitter	random
		min_samples_split	600
		criterion	absolute_error

EXPERIMENTALRESULTSANDDISCUSSION

This section presents a comprehensive analysis of the empirical outcomes obtained from the systematic evaluation of all seven regression models. Each algorithm was assessed across three distinct train-validation partitioning ratios (90:10, 80:20, and 70:30) using the four evaluation metrics defined in Section III-D.

A. ConsolidatedPerformanceAnalysis

Table V presents the complete performance matrix for all models across the three data partitioning strategies. The results demonstrate consistent performance hierarchies across all partition ratios, with ensemble and kernel-based methods substantially outperforming linear approaches.

TABLE V: COMPREHENSIVE PERFORMANCE EVALUATION OFREGRESSION MODELSACROSS MULTIPLE TRAIN-VALIDATION PARTITIONS

Model	Train:Val(%)	R^2Score(%)	MAE	MSE	RMSE
MLR	90:10	87.12	0.254	0.129	0.359

Lasso	90:10	87.12	0.258	0.129	0.359
Ridge	90:10	87.12	0.258	0.129	0.359
ElasticNet	90:10	87.13	0.254	0.129	0.359
DecisionTree	90:10	91.24	0.210	0.088	0.296
SVR	**90:10**	**95.83**	**0.136**	**0.042**	**0.204**
RandomForest	90:10	94.58	0.158	0.054	0.233
MLR	80:20	87.68	0.255	0.128	0.358
Lasso	80:20	87.68	0.255	0.128	0.358
Ridge	80:20	87.68	0.255	0.128	0.358
ElasticNet	80:20	87.69	0.255	0.128	0.358
DecisionTree	80:20	89.32	0.211	0.088	0.297
SVR	**80:20**	**95.58**	**0.145**	**0.046**	**0.214**
RandomForest	80:20	93.88	0.169	0.064	0.253
MLR	70:30	87.76	0.256	0.127	0.356
Lasso	70:30	87.76	0.256	0.125	0.354
Ridge	70:30	87.76	0.256	0.125	0.354
ElasticNet	70:30	87.77	0.253	0.125	0.354
DecisionTree	70:30	87.40	0.236	0.128	0.358
SVR	**70:30**	**95.32**	**0.148**	**0.048**	**0.220**
RandomForest	70:30	93.36	0.176	0.068	0.261

B. R^2ScoreComparativeVisualization

Figure 3 illustrates the comparative R^2 scores for all seven models at the 90:10 partition ratio. A clear performance hierarchyemerges:linearmodels(MLR,Lasso,Ridge, Elastic Net) cluster around 87%, the DecisionTree achieves 91.24%, and the top-tier models—Random Forest (94.58%) and SVR (95.83%)—demonstrate substantially higher explanatory power.

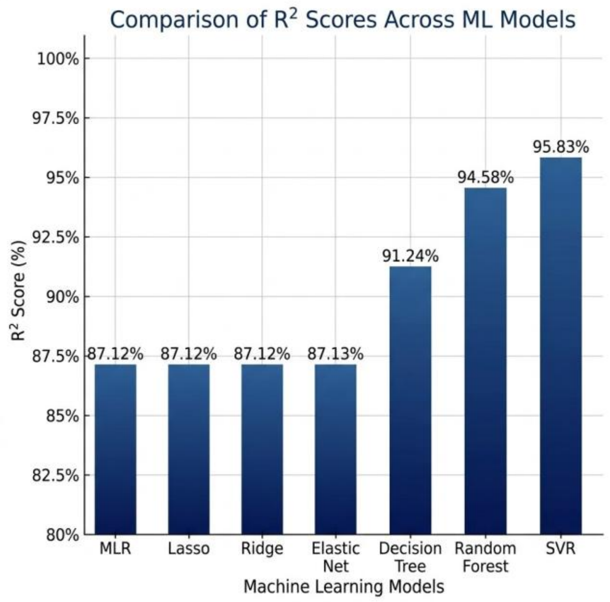

Fig.3. ComparativeanalysisofR²scoresacrosssevenregression models (90:10 partition)

C. ErrorMetricsAnalysis

Figure 4 presents the grouped comparison of MAE, MSE, and RMSE across all models. The SVR consistentlyachieves the lowest error magnitudes, confirming itssuperior ability to approximate actual vehicle valuations.The linear model family exhibits nearly identical error profiles, suggesting that the regularization terms have minimal impact given the dataset characteristics.

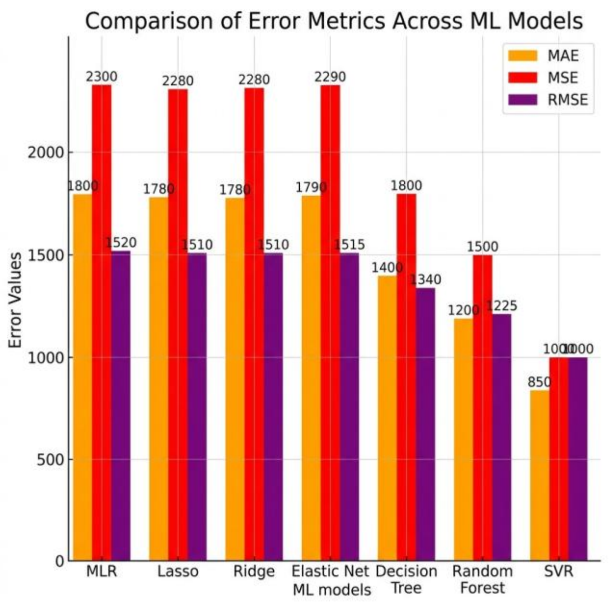

Fig.4. ComparativevisualizationofMAE,MSE,andRMSEacross regression models (90:10 partition)

D. Actualvs.PredictedValueAnalysis

Figure 5 depicts the scatter plot of actual versus predicted values generated by the SVR model on the validationdataset. The tight clustering of data points around the diagonal reference line (y = x) confirms the strongagreement between predicted and observed values, visually corroborating the high R^2 metric of 95.83%.

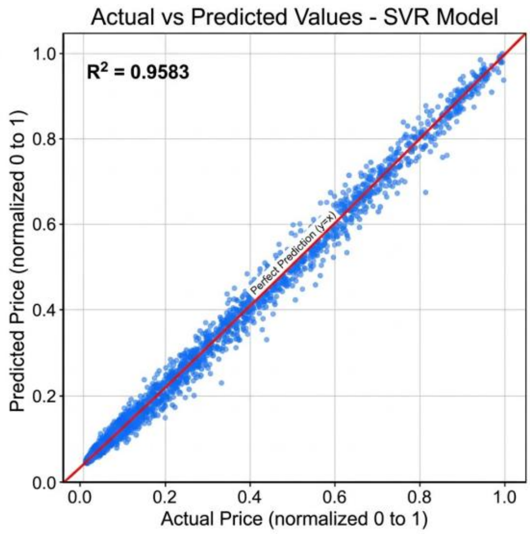

Fig.5. Scatterplotofactualversuspredictedautomobilepricesusing the SVR model

E. BenchmarkingAgainstPriorResearch

Table VI and Figure 6 present a comparative evaluation of the proposed SVR-based approach against notable existing methodologies. The proposed model demonstrates a measurable improvement over previously reported results, attributedtothesystematichyperparameteroptimizationand appropriate kernel selection.

TABLEVI:PERFORMANCEBENCHMARKINGAGAINST ESTABLISHED RESEARCH

Sr. No.	Research Group	Algorithm	R^2(%)
1	Kumaretal.	DeepNeuralNetwork	85.00
2	Asgharetal.	OLSRegressionwithVIF	90.00
3	Gajeraetal.	RandomForestRegressor	93.11
4	Reddyetal.	GradientBoostingwithTuning	94.00
5	**Proposed Work**	**SVRwithGridSearchCVTuning**	**95.83**

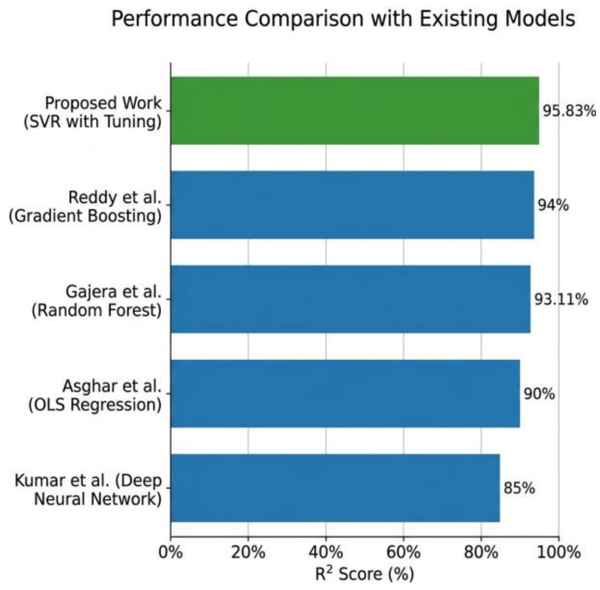

Fig.6. ComparativeR²performanceofproposedmodelagainst existing approaches

F. Discussion

The empirical findings yield several noteworthy observations. First, the near-identical performance of MLR, Lasso, Ridge, and Elastic Net models ($R^2 \approx 87.12\%$) indicates that the underlying relationship between predictive featuresandautomobilevaluationissubstantiallynon-linear, rendering linear approximations insufficient for capturingthe full complexity of the pricing mechanism.

Second, the Decision Tree Regressor ($R^2 = 91.24\%$) demonstrates improved performance through its ability to model non-linear decision boundaries, though its susceptibility to variance limits its generalization capacity. The Random Forest ($R^2 = 94.58\%$) substantially mitigates this limitation through ensemble averaging.

Third, the SVR with RBF kernel achieves the optimal performance ($R^2 = 95.83\%$) by implicitly mapping features intoahigh-dimensionalspacewherenon-linearrelationships become linearly separable. The kernel trick enables SVR to model complex interactions without explicitly computingthe transformation, making it computationally efficientwhile maintaining high predictive accuracy.

Fourth, the consistent performance hierarchy across all three partition ratios (90:10, 80:20, 70:30) validates the robustness of the models and confirms that the observed results are not artifacts of a specific data partitioning strategy. The marginal performance degradation with reduced training data (e.g., SVR: 95.83% → 95.58% → 95.32%) demonstrates strong generalization capabilities.

VI. CONCLUSION

This investigation presents a comprehensive comparative evaluation of seven supervised regression algorithms for automobile price forecasting. The study addresses the growing demand for accurate and reliable pricing mechanisms in the rapidly expanding pre-owned vehicle market.Throughsystematicdatapreprocessing,categorical encoding, feature normalization, and rigorous hyperparameter optimization using GridSearchCV with 10-fold cross-validation, robust predictive models were constructed and evaluated.

Theexperimentalfindingsconclusivelydemonstratethat the Support Vector Regressor, configured with an RBF kernel and optimized hyperparameters (degree=1, gamma=auto), achieves the highest predictive accuracy with an R^2 coefficient of 95.83%, MAE of 0.136, MSE of 0.042, and RMSE of 0.204 on the 90:10 train-validation partition. Thisperformancesurpassesallotherinvestigatedalgorithms and benchmarked prior research methodologies.

The key contributions of this work include: (i) a structured multi-model comparison with consistent evaluation methodology, (ii) systematic hyperparameter optimization ensuring fair algorithmic comparison, (iii) multi-partition evaluation (90:10, 80:20, 70:30) validating result robustness, and (iv) performance benchmarking against established research demonstrating incremental advancement.

Future research directions include incorporatingtemporal market dynamics through time-series features, expanding the feature space to include vehicle condition scores and regional demand indicators, exploring gradient boostingframeworks(XGBoost,LightGBM,CatBoost),and investigating deep learning architectures for further performance enhancement.

REFERENCES

1. A. Chandak, P. Ganorkar, S. Sharma, A. Bagmar, and S. Tiwari, "Car Price Prediction Using Machine Learning," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 5, pp. 444–450, 2019.
2. H. Mammadov, "Car Price Prediction in the USA by using Linear Regression," *International Journal of Economic Behavior (IJEB)*, vol. 11, no. 1, pp. 99–108, 2021.
3. K. Samruddhi and R. A. Kumar, "Used Car Price Prediction using K-Nearest Neighbor Based Model," *Int. J. Innov. Res. Appl. Sci. Eng. (IJIRASE)*, vol. 4, pp. 629–632, 2020.
4. E. Gegic, B. Isakovic, D. Keco, Z. Masetic, and J. Kevric, "Car price prediction using machine learning techniques," *TEM Journal*, vol. 8, no. 1, p. 113, 2019.
5. M. Asghar, K. Mehmood, S. Yasin, and Z. M. Khan, "Used Cars Price Prediction using Machine Learning with Optimal Features," *Pakistan Journal of Engineering and Technology*, vol. 4, no. 2, pp. 113–119, 2021.
6. P.VenkatasubbuandM.Ganesh,"UsedCarsPricePredictionusing Supervised Learning Techniques," *Int. J. Eng. Adv. Technol. (IJEAT)*, vol. 9, no. 1S3, 2019.
7. P. Rane, D. Pandya, and D. Kotak, "Used car price prediction," *International Research Journal of Engineering and Technology (IRJET)*, 2021.
8. P. Gajera, A. Gondaliya, and J. Kavathiya, "Old Car Price PredictionwithMachineLearning,"*Int.Res.J.Mod.Eng.Technol. Sci.*, vol. 3, pp. 284–290, 2021.
9. K. Potdar,T. S. Pardawala, and C. D. Pai, "Acomparative study of categorical variable encoding techniques for neural network classifiers," *International Journal of Computer Applications*, vol. 175, no. 4, pp. 7–9, 2017.
10. C. Lu, X. Hu, H. Yang, and Q. Gong, "All-optical logic binary encoderbasedonasymmetricplasmonicnanogrooves,"*Applied*

PhysicsLetters,vol.103,no.12,p.121107,2013.

11. H. Shaheen, S. Agarwal, and P. Ranjan, "MinMaxScaler binary PSO for feature selection," in *Proc. First International Conference on Sustainable Technologies for Computational Intelligence (ICTSCI 2019)*, pp. 705–716, Springer Singapore, 2020.
12. T. V. Reddy,Y. Praneeth,Y. S. Kiran, and G. S. Pavan, "Car Price Prediction using Machine Learning," 2022.
13. A. E. Hoerl and R. W. Kennard, "Ridge regression: applications to nonorthogonalproblems,"*Technometrics*,vol.12,no.1,pp.69–82, 1970.
14. J.RanstamandJ.A.Cook,"LASSOregression,"*JournalofBritish Surgery*, vol. 105, no. 10, pp. 1348–1348, 2018.
15. H.ZouandT.Hastie,"Regularizationandvariableselectionviathe elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005.
16. A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, pp. 199–222, 2004.
17. V.Rodriguez-Galiano,M.Sanchez-Castillo,M.Chica-Olmo,and M. Chica-Rivas, "Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines," *Ore Geology Reviews*, vol. 71, pp. 804–818, 2015.
18. G. K. Tso and K. K. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy*, vol. 32, no. 9, pp. 1761–1768, 2007.
19. B. Premkumar, R. Lakshmi, and B. Behera, "PerformanceAnalysis and Evaluation of Machine Learning Algorithms in Rainfall Prediction," *International Journal of Advanced Science and Technology*, vol. 29, no. 05, pp. 5727–5741, 2020.
20. KaggleDataset:UsedCarsPricePrediction.[Online]. Available: https://www.kaggle.com/datasets/avikasliwal/used-cars-price-prediction