International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 71-79
Published Online September 2022 in IJEAST (http://www.ijeast.com)

COMPARISON OF MACHINE LEARNING ALGORITHMS FOR PREDICTING DIAMOND PRICES BASED ON EXPLORATORY DATA ANALYSIS

Fadhil Muhammad Basysyar, Gifthera Dwilestari
Information System
STMIK IKMI, Cirebon, West Java, Indonesia

Abstract—**Diamonds are a unique commodity whose socially generated notions significantly influence perceived value. To study how a diamond's physical attributes might predict its price, a massive dataset of loose diamonds scraped from an online diamond store is subjected to data mining, which reveals that diamond weight, color, and clarity are the most influential determinants of diamond pricing. Therefore, submit a proposal for an Exploratory Data Analysis that includes a component that analyses various parts of news articles using LASSO Regression, ElasticNet Regression, and Random Forest Regression. This system is trained on past data to forecast diamond prices while retaining an easily interpretable trading approach concerning rule complexity. The suggested strategy beats cutting-edge methods for prediction accuracy and interpretability, such as extreme learning machines using deep learning. Our data indicate that the news impact factor is crucial for forecasting. Demonstrate that the suggested system outperforms the average yearly return while offering a set of language trading rules that are interpretable. This has substantial repercussions for investors. A significant degree of subjectivity in diamond pricing may result from diamond dealers' price concealment techniques.**

Keywords—**Exploratory Data Analysis, LASSO Regression, ElasticNet Regression, and Random Forest Regression.**

I. INTRODUCTION

Diamonds are a unique consumer commodity, and one has no apparent benefit. Although diamonds are the most complex jewel on Earth, 58 times harder than anything else[1]–[4], they are rarely used for this purpose. One of the most obvious motivations for buying or wearing diamonds is the perception of their rarity and expense. Although diamonds have been prized for millennia, their consumption altered substantially in the 20th century when diamond engagement rings became fashionable. De Beers, the most prominent operator in the diamond business, has used the tagline "a diamond is forever" since 1948 [5] to promote such contributions. The United States accounted for 39 billion dollars of the total global diamond jewelry sales, reaching $79 billion [6]. The worldwide demand for polished loose diamonds amounted to $25 billion [7].

Traditionally, diamonds were sold at jewelry stores. With the expansion of e-commerce, more diamonds are being offered [8]–[11]online, providing a more significant consumer base. Even though the price paid for a diamond can affect its perceived worth , diamonds have well-defined physical attributes such as weight, cut, and color. In addition to boosting new markets for diamonds, the rise of e-commerce is anticipated to reduce customer search costs and make it simpler for consumers to compare diamond pricing and physical attributes among shops [12], [13].

In the past decade, statistical and artificial intelligence techniques for forecasting the price of diamonds and other jewels have increased in popularity [14]. Diamond's unique properties, such as its function as a financial asset, store of monetary value, and supply buildup, make it challenging to anticipate diamond prices, as real-world diamond price data contradict typical statistical assumptions [15]–[17]. Thus, artificial intelligence prediction models such as Neural Networks [18] and decision trees [19] surpass classic statistical models in terms of accuracy. Existing prediction models based on artificial intelligence have significant drawbacks, notwithstanding their accuracy. First, these models' lack of openness and interpretability prohibits investors from employing them as decision-making aids [20], [21]. Due to the numerous uses of diamonds and the extensive linkages between the diamond market and other financial and commodity markets [22], [23], the diamond price time series also includes a high amount of intrinsic uncertainty. The inability of present forecasting models to appropriately account for the influence of news events on diamond prices is another restriction. Existing work focuses on the effects of news mood on financial markets, whereas commodities markets have received less attention [24].

The current work provides an interpretable, automated approach for predicting the price of diamonds using the popularity and strength of sentiments in financial news. The proposed prediction system consists of many components [25]based on Exploratory Data Analysis, including a component that analyses various features of news articles using LASSO Regression, ElasticNet Regression, and Random Forest Regression. This study examines the advantages of this strategy over others, such as its excellent computing efficiency and the linguistic quality of the induced rules [26]. The proposed prediction method is highly competitive with cutting-edge technology in artificial intelligence. Nonetheless, it also supplies investors with a decision-making aid constituted of linguistic trading principles that may be understood. Profit-wise, the trading strategy based on the proposed prediction algorithm outperforms methods. Consequently, our prediction technology may greatly minimize price volatility uncertainty and facilitate future investment decisions.

The paper's organization is as follows: Provide an overview of prior studies on consumer search and pricing variation. Examine the critical diamond physical characteristics that are known to affect diamond pricing. Then, review the research on price opacity or merchant strategies to increase consumer search costs. Explain the dataset, the technique, and the data mining insights by highlighting the results, contributions, and future avenues for study.

II.LITERATURE REVIEW

In today's world, it is more challenging to predict the price of certain commodities, such as diamonds. Ultimately, diamond is precious and is held in the strategic reserves of most central banks [27], [28]. Mining companies are affected by the difficulty of pricing forecasting since they must include future prices in their operations. In contrast, pessimistic projections may restrict, terminate, temporarily halt, or reduce mining activities. On the other hand, positive occurrences and anticipations increase or prolong mining activity relative to an organization's initial estimations. Consequently, an accurate forecast model is crucial for both the investor community and the investment choices of mining businesses. Previous writers have underlined [29] that it is crucial for commodities markets and the global economy to precisely estimate diamond price fluctuations.

The price fluctuations of diamonds have been modeled using traditional mathematical and statistical time series prediction techniques. A model with mixed data sampling was utilized to illustrate how global policy unpredictability may aid in predicting future volatility [30]. The mathematical model explored the volatility spillovers between the stock market and the price [31], [32]. Significant cross-effects of volatility and return were identified, and diamond was identified as a stock hedge. A dynamic conditional correlation model [33] was also utilized to determine a diamond's safe-haven status, supporting the hypothesis that there are substantial linkages between the equity and commodity markets. In models characterizing conditional pricing volatility, short- and long-term volatility components were recorded [34]–[36].

Using a modified model, the link between the price of gold and its determinants, such as oil price, was illustrated [37]. In terms of forecasting mistakes, this method fared better than the model. Although previous research has been utilized to forecast prices [38], [39], adjustments such as multivariate empirical mode decomposition to detect noise sources led to significant gains [40]–[42]. Granger's causality method was used to search for evidence of bidirectional causality pricing [43]. Prior empirical data has shown that special prices react nonlinearly to price changes and that bidirectional causality pricing exists [44], [45]. This link was confirmed further [46]–[48].

Similarly, it has been demonstrated that the stock market and inflation hedge directly affect the price [49]. These mathematical and statistical models provided vital insight into the elements that influence price fluctuations. However, assumptions such as homoscedasticity and stationarity of time series restrict the effectiveness of these models' difficulty in applying these constraints to data.

In anticipating prices, artificial intelligence systems have beaten mathematical and statistical methodologies and an adaptive inference system[50], [51]. Predicting price return volatility with a hybrid model significantly decreased error[52], [53]. The whale optimization strategy was employed to overcome the poor prediction performance of typical gradient-based training methods because of their convergence to local minima of the error function [54]. The resultant model outperformed traditional models that were trained using alternative evolutionary procedures. A recent study [55], [56] has revealed that long-term and short-term memory is superior due to its ability to acquire complicated, high-level temporal features from time-series data. Prior research has created strategies for forecasting metal prices. Unpredictable price fluctuations are modeled using a graphical prediction approach [57], [58]. According to empirical evidence, this model outperformed the competition. Several artificial intelligence approaches have been used to estimate costs, including an algorithmically optimized time series function, a combination of a price volatility network, and a collection of inference system models.

II. METHODOLOGY

A. Data Preprocessing

The objective is to estimate the price of a diamond based on its many attributes. There are roughly 54,000 diamonds, each with 11 distinguishing features. The average diamond price was then predicted based on these characteristics, which acted as dataset features. The features of the Carat diamond are seen in Fig.1: The diamond's weight, which is comparable to 200 milligrams, should be an excellent signal. Cut: The quality of the cut. Color: The diamond's color, from worst to best (J to D) Clarity relates to the diamond's transparency. Percentage of total depth relative to x and y, representing depth. Table: Top

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 71-79
Published Online September 2022 in IJEAST (http://www.ijeast.com)

of diamond width relative to its widest point, price: in US dollars, and dimensions: x, y, and z.

```
RangeIndex: 53940 entries, 0 to 53939
Data columns (total 11 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Unnamed: 0  53940 non-null  int64
 1   carat       53940 non-null  float64
 2   cut         53940 non-null  object
 3   color       53940 non-null  object
 4   clarity     53940 non-null  object
 5   depth       53940 non-null  float64
 6   table       53940 non-null  float64
 7   price       53940 non-null  int64
 8   x           53940 non-null  float64
 9   y           53940 non-null  float64
 10  z           53940 non-null  float64
```

Fig. 1. Diamond Characteristics

Before starting the analysis, it is crucial to delete or fill null entries and change the data type of misspecified columns. In feature engineering, outliers in the dataset were examined. Interquartile Range can be used to detect an outlier x if the following conditions are met:

$$x < Q1 - 1.5 * IQR \quad OR \quad Q3 + 1.5 * IQR < x \qquad (1)$$

where:
Q1 = 25th percentiles
Q3 = 75th percentiles
IQR = Q3 − Q1

After applying the equation (1) to each dataset's column, the resulting dataset had 53,920 data points with 11 attributes. Table 1 provides a detailed description of the Interquartile Range, while Fig. 2 and Fig. 3 depict the Quantitative Distribution. The tails of the depth, table, y, and z distributions are long. There is a value in z that looks like an error or severe outlier, and exploring other outliers is possible. x, y, and z should not have the value 0 since it is illogical. When any of these variables is 0, remove or impute all rows. Due to the absence of context, avoid removing outliers wherever possible. This z value may be accurate. Next, examine the category variables for any apparent issues.

Table - 1 Interquartile Range

	carat	depth	table	price	x	y	z
count	53940.000000	53940.000000	53940.000000	53940.000000	53940.000000	53940.000000	53940.000000
mean	0.797940	61.749405	57.457184	3932.799722	5.731157	5.734526	3.538734
std	0.474011	1.432621	2.234491	3989.439738	1.121761	1.142135	0.705699
min	0.200000	43.000000	43.000000	326.000000	0.000000	0.000000	0.000000
25%	0.400000	61.000000	56.000000	950.000000	4.710000	4.720000	2.910000
50%	0.700000	61.800000	57.000000	2401.000000	5.700000	5.710000	3.530000
75%	1.040000	62.500000	59.000000	5324.250000	6.540000	6.540000	4.040000
max	5.010000	79.000000	95.000000	18823.000000	10.740000	58.900000	31.800000

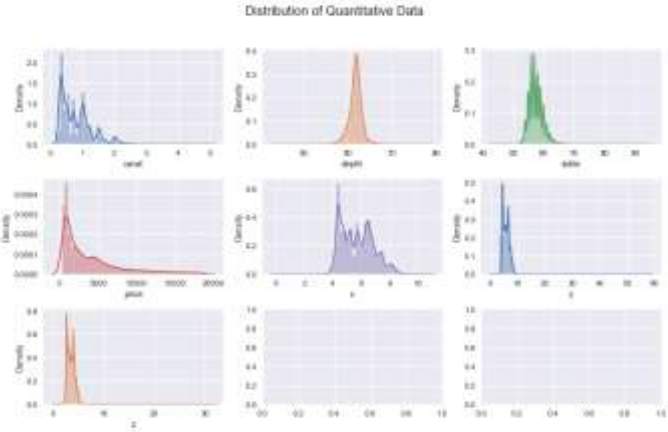

Fig. 2. Distibution of Quantitative

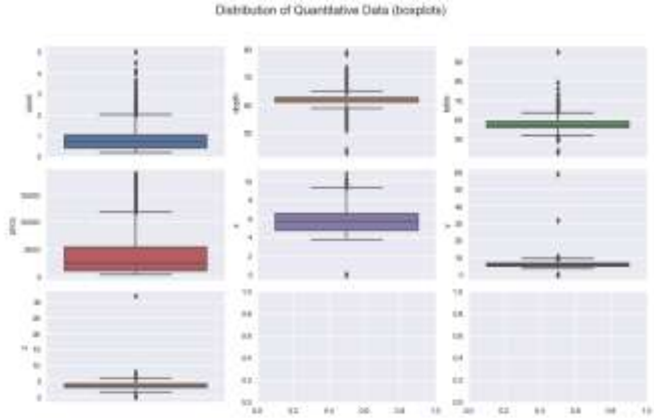

Fig. 3. Distibution of Quantitative Boxplot

B. Data Analysis

Exploratory data analysis is required prior to constructing a prediction model and enables the identification of implicit data patterns, which aids in selecting suitable machine learning algorithms. Determine which predictors are correlated with the price of a diamond, whether high-quality diamonds are more expensive than low-quality diamonds if there are any direct interactions between a categorical predictor, a numerical predictor, and the response if the dataset exhibits collinearity or multicollinearity, and if there are any apparent outliers that warrant further investigation.

Figure 4 demonstrates that price is biased to the right. Consequently, log transformation is required for more precise forecasting tests to discover the optimal combination, as there are no obvious clues in this circumstance. In addition, there is an apparent anomaly that warrants more research. In addition, x is related to the diamond's carat, which might lead to a collinearity issue. Check the correlation matrix to be specific.

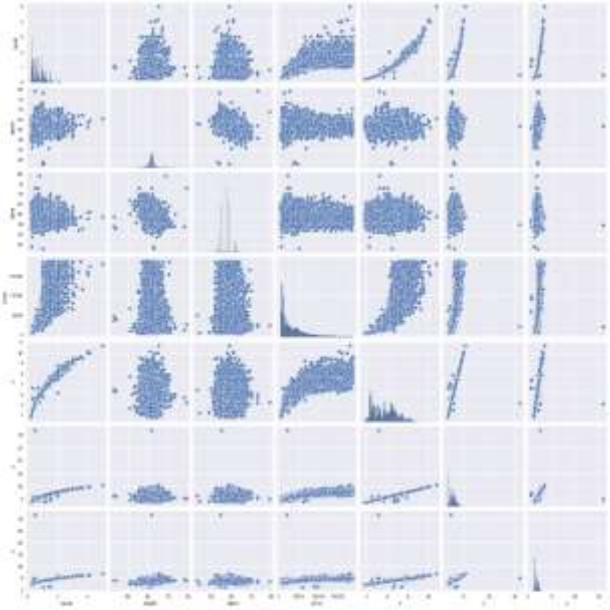

Fig. 4. Experiment to Find Best Combinations

Examine an uncommon outlier. An attempt was made to disregard the outlier, but it looks to be an obvious mistake. Examine the peculiar anomaly in the y and z charts. The two outliers may have a mean of 3.18 instead of 31.8, which is an assumption but appears reasonable. These two numbers out of 50,000 are anomalous and will undoubtedly impact regression models that lack robustness against outliers. Iterative methods such as random forest and gradient boosting can handle them, but they presume that they are mistakes and have a low probability of occurring naturally in the actual world. Note that this is not necessarily the ideal approach to dealing with outliers; imagine there was a type of observation that was not recorded because someone misplaced all the diamonds of this type, as seen in Fig. 5.

```
count    43136.000000
mean         3.539305
std          0.704325
min          1.070000
25%          2.910000
50%          3.520000
75%          4.030000
max         31.800000
Name: z, dtype: float64
```

Fig. 5. Two Outliers

There is evident collinearity when two outliers are excluded. x, z, and z are interconnected and should be integrated, or just one should be utilized to estimate the cost. Carat is also strongly associated with x, y, and z. Carat is a function of dimensions and a density coefficient, but this is not necessarily the case because carat is most frequently a unit of weight (source). As seen in Fig. 6 and Fig. 7, delete x, y, and z.

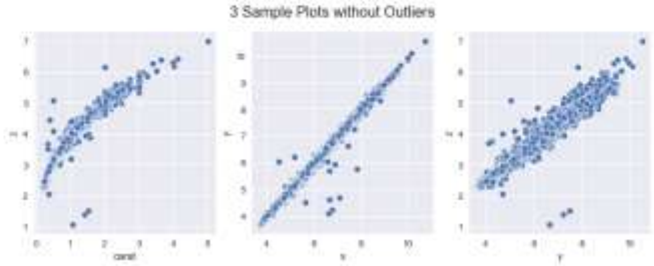

Fig. 6. Sample Plots Without Outliers

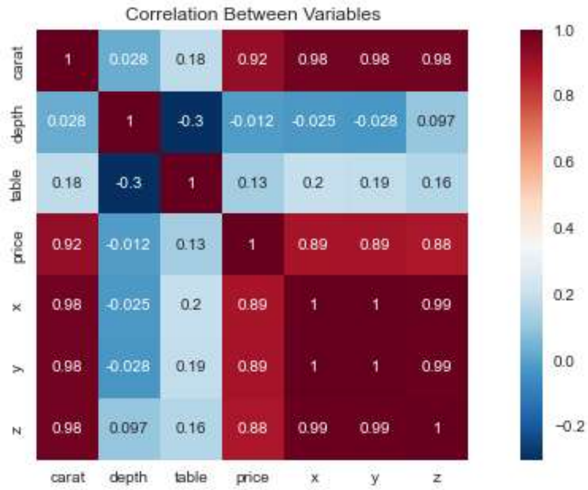

Fig. 7. Correlation Between Variables

Examine the link between a diamond's quality and price using categorical variables and a combination of continuous and categorical parameters. When the categories are arranged, a correlation exists between carat, price, and color. The same holds for carat, price, and clarity. However, the interaction between carat and cut is not as evident, as seen in Fig. 8, Fig. 9, and Fig. 10.

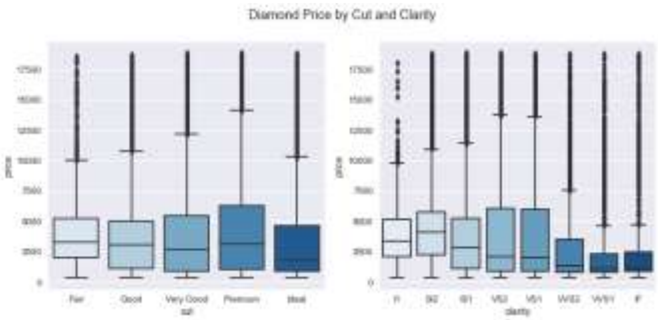

Fig. 8. Diamond Price by Cut and Clarity

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 71-79
Published Online September 2022 in IJEAST (http://www.ijeast.com)

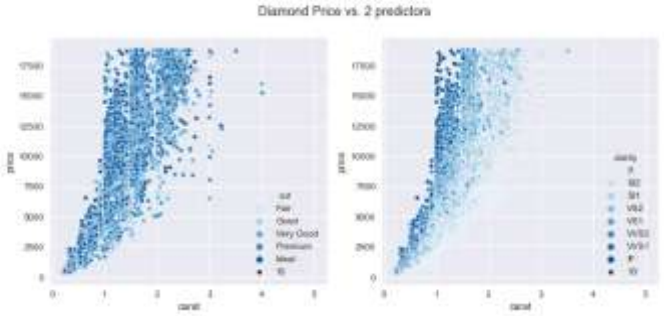

Fig. 9. Diamond Price vs. 2 Predictors

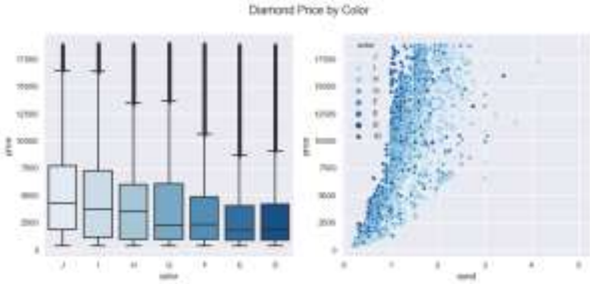

Fig. 10. Diamond Price by Color

Upon examination of the data summary, the answer looked to be skewed. Although not essential, it is possible to visualize the effect of various transformations on the result, such as square root, log, and cube root. Experiment with this throughout our modeling phase Fig. 11 depicts pricing modifications.

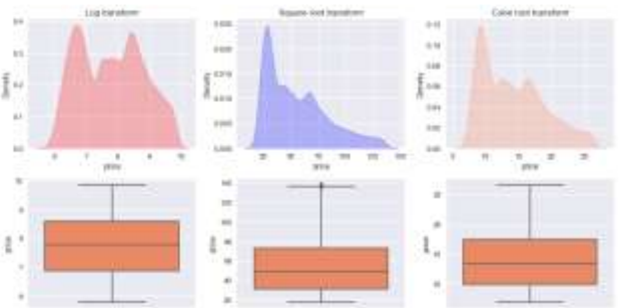
Fig. 11. Transformations on Price

IV.RESULT AND DISCUSSION

A. Model Selection and Model Build

Consider the following models of regression: Ridge Regression, LASSO Regression, ElasticNet, Random Forest Regression, XGBoost, Support Vector Regression, and a stacked ensemble were all investigated, however SVR was too slow owing to a huge number of features, and basic models performed well, indicating that more sophisticated models were unnecessary. Since the training and test sets have already been split, we should begin by one-hot encoding our category variables before normalizing the numerical variables. Provide both an interpretable and a model that is likely to outperform interpretability.

Mean Absolute Error is a measurement of error. Mean Squared Error is the model producing substantial errors, while R2 Score is the model's goodness-of-fit. Due to the fact that all models have the same amount of predictors p, the R2adj score is superfluous and the outlier is eliminated from our X train. Allows models to acquire more relevant information. Observations with 0x, y, or z values were eliminated using Missing Values. There are no missing values remaining.
The categorical variables are ordinal rather than merely categorical, necessitating the usage of encoding. An ordered representation of the categories is employed, which is then translated to matching numbers for the additional columns. Use MinMax scaling after encoding, with the possibility of experimenting with Standard Scaler if all predictors are somewhat typical. Since there are few outliers, strong transformers are unnecessary.

B. Ridge Regression

As seen in Fig. 12, Ridge Regression creates and fits a list of error distributions for each of 13 characteristics, totaling 53,920 data points. Fig. 9 depicts the actual and expected pricing parameter values for negative mean absolute error and alpha.

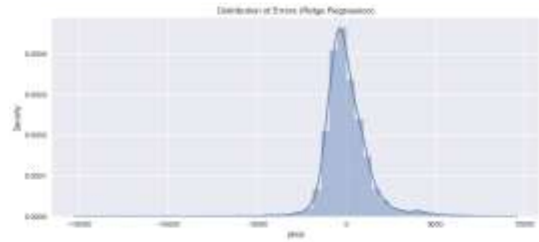
Fig. 12. Distribution of Errors


Fig. 13. Actual vs. Predicted Ridge Regression

There are still other factors to consider. This research attempts to construct a model capable of predicting the price of a diamond in a specific environment based on a set of characteristics. The model will be used to evaluate the strength of the correlations between the response and the predictors, which is a crucial objective when constructing one.

C. LASSO Regression

Therefore, an intelligent approach to feature reduction that does not affect model performance must be developed. The Lasso algorithm should perform better when just a small number of the predictors used to construct our model have a meaningful impact on the response variable. As a result, it also serves as a method for feature selection, eliminating irrelevant variables: Fig. 14 depicts the error distribution of the Lasso Regression.

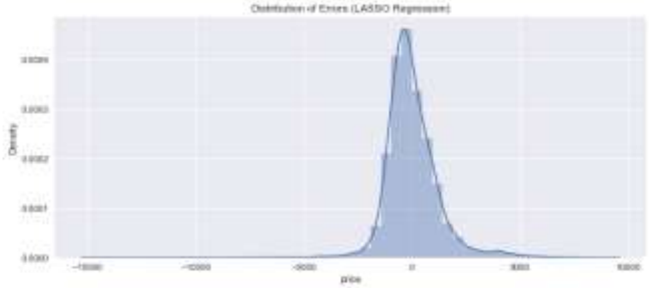

Fig. 14. Lasso Regression Distribution of Errors

Lasso performs better than Ridge for predicting unknown data. The regression loss is about equivalent. Instead of utilizing the optimal alpha value for feature selection, slightly raise it. It also serves as a tool for selecting features and minimizing superfluous variables. Similar to the ridge regression model, it appears that the model under predicts the price of $10,000 diamonds, as seen by the histogram (mean is less than zero) and scatter plot (the trend is a curve below the y=x line).

Fig. 15. Actual vs. Predicted LASSO Regression

D. ElasticNet Regression

Therefore, an intelligent approach to feature reduction that does not affect model performance must be developed. The Lasso algorithm should perform better when just a small number of the predictors used to construct our model have a meaningful impact on the response variable. As a result, it also serves as a tool for feature selection, removing irrelevant factors. Lasso performs better than Ridge for predicting unknown data. The regression loss is about equivalent. Instead of utilizing the optimal alpha value for feature selection, slightly raise it. It also serves as a tool for selecting features and minimizing superfluous variables. Both models provide comparable results. Therefore let us test whether integrating both kinds of penalties (L1 and L2) increases prediction accuracy through a series of tests with l1 ratio values of 0.5, 0.3, 0.8, and 0.9. Observe that having an L1 penalty decreases errors and improves R2. Having a minor L2 penalty, however, lowered our mistake rate.

E. Random Forest Regression

Sklearn offers the Random Forest Classifier class for constructing Random Forest Classifier's n estimators parameter. While increasing the number of trees in a random forest improves accuracy, it also lengthens the model's total training time. Additionally, the class has the bootstrap parameter set to True. In the case of random forest subsets, however, only a restricted number of attributes will be employed to give variety to the trees; efficiency is increased by repeatedly iterating the model and adding a few choices when initializing the Random Forest Classifier.

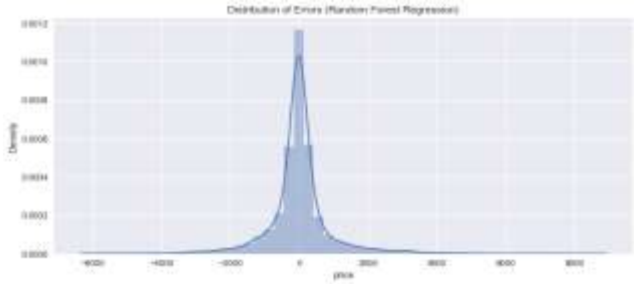

Fig. 16. Distribution of Errors Random Forest Regression

Use random forest regression to assess a nonlinear model based on the data. Initially, Support Vector Regression was examined. However, due to several observations, it was prolonged (see the SVR documentation on sklearn for other approaches). First, train a model with one thousand estimators, and if the model looks improvable, perform a grid search. With max depth = 2 and max depth = 5, the model was modified, with max depth = 5 greatly rising. Based on our three error measurements, this model is superior.

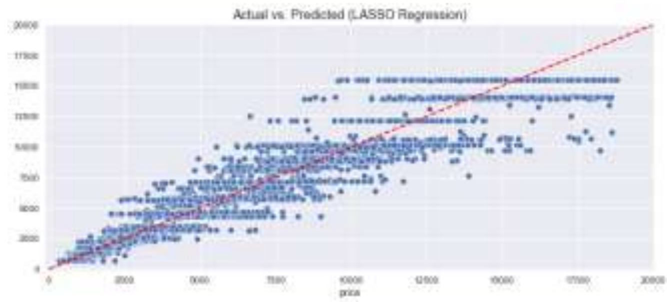

Fig. 17. Actual vs. Predicted

V.CONCLUSION

Predicting diamond prices as a function of diamond physical properties using predictive data mining techniques on a dataset of 53,920 diamonds scraped from a leading online diamond retailer, drawing on theoretical and empirical research on the association between consumer search costs and price dispersion, is investigated. Random Forest $538.72 for Regression, $854.20 for ElasticNet (with 0.9 L1 penalty), $860.88 for Ridge, and $861.01 for LASSO. The impact of several diamond physical characteristics on price finds that diamond weight and associated dimensions (length, width, and height) play the most important effect, followed by color, clarity, and shape. The first data mining findings had substantial inaccuracies, reducing the diamond weight range to 0.2 to 2.5 carats. This range was selected based on empirical data indicating that people purchase these sorts of diamonds, most commonly online. The artificial neural network model yielded the best mean absolute percent error findings. In other words, an advanced modeling computer could not anticipate diamond values based on their physical characteristics.

Algorithms' extremely high degree of inaccuracy shows a significant degree of subjectivity and, hence, price dispersion in diamond pricing. The investigation of price distributions for diamonds weighing half a carat and one carat, which are popular with buyers, revealed more evidence of price dispersion. Our data indicate that price dispersion increases dramatically for diamonds at or beyond these critical thresholds. The standard variation of diamond prices exceeds 47 percent of the average price for 1.04-carat diamonds; this is especially significant given that our data is taken from a single store whose website allows users to compare gems without leaving the site.

Lastly, it is crucial to note that no research is devoid of constraints. An experimental study was done using a dataset from a single online diamond retailer. The ideas may be inapplicable in specific contexts. For instance, the physical purchasing experience for diamonds will be drastically different. Anecdotal information suggests that the diamond industry is relatively opaque. The scope of our research is limited to the information acquired on the retailer's website. These data give no information on why customers purchase diamonds. For instance, customers may purchase diamonds as an investment, which may be impacted by reasons other than those investigated in our study.

VI. REFERENCE

[1]. Smith E. M., Ni P., Shirey S. B., Richardson S. H., Wang W., and Shahar A. (2021). Heavy iron in large gem diamonds traces deep subduction of serpentinized ocean floor. Sci Adv.doi: 10.1126/sciadv.abe9773.

[2]. Ma Y. et al.(2021). The catalysts of synthetic gem-quality HPHT diamonds and their impact on diamond quality: A case study of synthetic diamonds from three Chinese companies.Zhongshan DaxueXuebao/Acta ScientiarumNatralium Universitatis Sunyatseni.doi: 10.13471/j. cnki.acta.snus.2020.11.23.2020D071.

[3]. Wang Z.et al. (2020). Synthesis and characterization of gem diamond single crystals in Fe-C system under high temperature and high pressure.J Cryst Growth.doi: 10.1016/j.jcrysgro.2019.125371.

[4]. Zhang K., Tian Y., Liu K., Zhang J., and Wang T. (2020). Synthesis of gem grade diamond by temperature gradient method using high temperature sintered alumina ceramics.JingangshiyuMoliaoMojuGongcheng/Diamond and Abrasives Engineering.doi: 10.13394/j.cnki.jgszz.2020.6.0002.

[5]. Laurin C., KalavrouziotisD., and Mohammadi S. (2021). Commentary: Diamonds are forever: Not so for transcatheter aortic valve replacement.Journal of Thoracic and Cardiovascular Surgery.doi: 10.1016/j.jtcvs.2021.09.050.

[6]. World Diamond Council. (2017). The Diamond Industry Fact Sheet.DiamondFacts.

[7]. Thomsen L. and Hess M. (2022). Dialectics of Association and Dissociation: Spaces of Valuation, Trade and Retail in the Gemstone and Jewelry Sector.Econ Geogr.doi: 10.1080/00130095.2021.1989302.

[8]. MamonovS. and TriantoroT. (2018). Subjectivity of diamond prices in online retail: Insights from a data mining study.Journal of Theoretical and Applied Electronic Commerce Research.doi: 10.4067/S0718-18762018000200103.

[9]. Wu C. and CosgunerK. (2020). Profiting from the decoy effect: A case study of an online diamond retailer.Marketing Science. doi: 10.1287/mksc.2020.1231.

[10]. Wolff F. C. (2016). Bargaining powers of buyers and sellers on the online diamond market: a double perspective non-parametric analysis.Ann Oper Res. doi: 10.1007/s10479-016-2160-1.

[11]. GudhlangaJ. and Spiegel S. J. (2021). Gendered social media communication around mining: patriarchy, diamonds, and seeking feminist solidarity online.Gend Dev.doi: 10.1080/13552074.2021.1982179.

[12]. SharifeK. and BrackingS. (2016). Diamond pricing and valuation in South Africa’s extractive political economy.Rev Afr Polit Econ. doi: 10.1080/03056244.2016.1177504.

[13]. Chu S. (2001). Pricing the C’s of Diamond Stones.Journal of Statistics Education.doi: 10.1080/10691898.2001.11910659.

[14]. Moffat I. U. and Akpan E. A. (2019). Selection of Heteroscedastic Models: A Time Series Forecasting Approach.Appl Math (Irvine).doi: 10.4236/am.2019.105024.

[15]. Mills T. C. (2004). Statistical analysis of daily gold price data.Physica A: Statistical Mechanics and its Applications.doi: 10.1016/j.physa.2004.03.003.

[16]. Hui K. and Lei S. (2022). The Design of the Benchmark Land Price Release and Management Information System Based on Webgis.Frontiers in Business, Economics and Management.doi: 10.54097/fbem.v3i1.230.

[17]. Lee M. C., Chang J. W., Yeh S. C., Chia T. L., Liao J. S., and Chen X. M. (2022). Applying attention-based BiLSTM and technical indicators in the design and performance analysis of stock trading strategies.Neural Comput Appl.doi: 10.1007/s00521-021-06828-4.

[18]. HaykinS. (2008).Neural Networks and Learning Machines. doi: 978-0131471399.

[19]. BulacC. and BulacA. (2016). Decision Trees. in Advanced Solutions in Power Systems: HVDC, FACTS, and AI Techniques. doi: 10.1002/9781119175391.ch18.

[20]. Tan A., Copley J., and Fleming J. (2021). Decision-making aids for upper limb interventions in neurological rehabilitation: a scoping review.Disability and Rehabilitation.doi: 10.1080/09638288.2021.1924881.

[21]. Lim A. H., StreeperN. M., Best S. L., PennistonK. L., and Nakada S. Y. (2017). Clinical use of patient decision-making aids for stone patients.Canadian Journal of Urology. vol. 24, no. 4.

[22]. Shah A. A. and Dar A. B. (2022). Asymmetric, time and frequency-based spillover transmission in financial and commodity markets.J Econ Asymmetries.doi: 10.1016/j.jeca.2022.e00241.

[23]. UromC., Ndubuisi G., and GuesmiK. (2022). How do financial and commodity markets volatility react to real economic activity?.Financ Res Lett.doi: 10.1016/j.frl.2022.102733.

[24]. RiméléA., DimitrakopoulosR., and Gamache M. (2020). A dynamic stochastic programming approach for open-pit mine planning with geological and commodity price uncertainty.Resources Policy.doi: 10.1016/j.resourpol.2019.101570.

[25]. Ho W. K. O., Tang B. S., and Wong S. W. (2021). Predicting property prices with machine learning algorithms.Journal of Property Research.doi: 10.1080/09599916.2020.1832558.

[26]. Chen C., Kim J. B., Wei M.Zhang. (2019). Linguistic Information Quality in Customers' Forward-Looking Disclosures and Suppliers' Investment Decisions.Contemporary Accounting Research.doi: 10.1111/1911-3846.12471.

[27]. Tule M. K., Salisu A. A., and ChiemekeC. C. (2019). Can agricultural commodity prices predict Nigeria's inflation?.Journal of Commodity Markets.doi: 10.1016/j.jcomm.2019.02.002.

[28]. Ge Y. and Tang K. (2020). Commodity prices and GDP growth.International Review of Financial Analysis.doi: 10.1016/j.irfa.2020.101512.

[29]. Chen Z., Goh H. S., Sin K. L., Lim K., N. Chung K. H., and Liew X. Y. (2021). Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques.Advances in Science, Technology and Engineering Systems Journal.doi: 10.25046/aj060442.

[30]. Kim J. Park J. (2020). Predictability of OTC option volatility for future stock volatility.Sustainability (Switzerland).doi: 10.3390/su12125200.

[31]. Tang Y., Xiao X.Wahab M. I. M., and Ma F. (2021). The role of oil futures intraday information on predicting US stock market volatility.Journal of Management Science and Engineering.doi: 10.1016/j.jmse.2020.10.004.

[32]. KanamuraT. (2022). A model of price correlations between clean energy indices and energy commodities.Journal of Sustainable Finance and Investment.doi: 10.1080/20430795.2020.1753434.

[33]. JaśkowskiP. and CzarnigowskaA. (2019). Contractor's bid pricing strategy: A model with correlation among competitors' prices.Open Engineering.doi: 10.1515/eng-2019-0021.

[34]. Chen R. and Xu J. (2019). Forecasting volatility and correlation between oil and gold prices using a novel multivariate GAS model.Energy Econ.doi: 10.1016/j.eneco.2018.11.011.

[35]. Liang Y. and Xu C. (2020). An efficient conditional Monte Carlo method for European option pricing with stochastic volatility and stochastic interest rate.Int J Comput Math.doi: 10.1080/00207160.2019.1584671.

[36]. RupandeL., MugutoH. T., and MuzindutsiP. F. (2019). Investor sentiment and stock return volatility: Evidence from the Johannesburg Stock Exchange.Cogent Economics and Finance.doi: 10.1080/23322039.2019.1600233.

[37]. PriyadiI., SantonyJ., and Na'amJ. (2019). Data Mining Predictive Modeling for Prediction of Gold Prices Based on Dollar Exchange Rates, Bi Rates and World Crude Oil Prices.Indonesian Journal of Artificial Intelligence and Data Mining.doi: 10.24014/ijaidm.v2i2.6864.

[38]. Hajek P. and Novotny J. (2022). Fuzzy Rule-Based Prediction of Gold Prices using News Affect.Expert Syst Appl. doi: 10.1016/j.eswa.2021.116487.

[39]. SuryanaY. and Sen T. W. (2021). The Prediction of Gold Price Movement by Comparing Naive Bayes, Support Vector Machine, and K-NN.JISA(JurnalInformatika dan Sains).doi: 10.31326/jisa.v4i2.922.

[40]. Makala D. and Li Z. (2021). Prediction of gold price with ARIMA and SVM. in Journal of Physics: Conference Series. doi: 10.1088/1742-6596/1767/1/012022.

[41]. Zhu Y. and Zhang C. (2018). Gold Price Prediction Based on PCA-GA-BP Neural Network.Journal of Computer and Communications.doi: 10.4236/jcc.2018.67003.

[42]. Brabenec T., SulerP., HorakJ., and PetrasM. (2020). Prediction of the future development of gold price.Acta MontanisticaSlovaca.doi: 10.46544/AMS.v25i2.11.

[43]. Khan I., Hou F., and Le H. P. (2021). The impact of natural resources, energy consumption, and population growth on environmental quality: Fresh evidence from the United States of America.Science of the Total Environment.doi: 10.1016/j.scitotenv.2020.142222.

[44]. AkinloA. E. (2012). How Important is Oil in Nigeria’s Economic Growth?.J Sustain Dev.doi: 10.5539/jsd.v5n4p165.

[45]. Nugroho R. E. (2019). Domestic Factors That Affect The Price of Styrene Butadiene Latex in Indonesia.Jurnal Ekonomi Pembangunan: Kajian MasalahEkonomi dan Pembangunan.doi: 10.23917/jep.v20i1.7097.

[46]. Samuel U. E., Rosemary I. H., InimV., EdedemA. J., and NdubuakuV. (2021). Energy consumption and sectorial value addition on economic growth in Nigeria.Universal Journal of Accounting and Finance.doi: 10.13189/ujaf.2021.090108.

[47]. PicksonR. B., He G., and Boateng E. (2022). Impacts of climate change on rice production: evidence from 30 Chinese provinces.Environ Dev Sustain.doi: 10.1007/s10668-021-01594-8.

[48]. Wang X., Khurshid A., Qayyum S., and Calin A. C. (2022). The role of green innovations, environmental policies and carbon taxes in achieving the sustainable development goals of carbon neutrality.Environmental Science and Pollution Research.doi: 10.1007/s11356-021-16208-z.

[49]. KumbureM. M., LohrmannC., LuukkaP., and Porras J. (2022). Machine learning techniques and data for stock market forecasting: A literature review.Expert Systems with Applications. doi: 10.1016/j.eswa.2022.116659.

[50]. Zhan M. F., Cai Z. W., Fang Y., and Lin M. (2022). Recent advances in statistical methodologies in evaluating program for high-dimensional data.Appl Math (Irvine).doi: 10.1007/s11766-022-4489-3.

[51]. PatiñoI. D. and IsazaC. A. (2022). Mori-Tanaka-based statistical methodology to compute the effective Young modulus of polymer matrix nano-composites considering the experimental quantification of nanotubes dispersion and alignment degree.Engineering Solid Mechanics.doi: 10.5267/j.esm.2021.9.002.

[52]. VeitA. and BelongieS. (2020). Convolutional Networks with Adaptive Inference Graphs.Int J Comput Vis.doi: 10.1007/s11263-019-01190-4.

[53]. HollingC. S. and Allen C. R. (2002). Adaptive inference for distinguishing credible from incredible patterns in nature.Ecosystems.doi: 10.1007/s10021-001-0076-2.

[54]. Valenstein-MahH. et al.(2020). Effectiveness of training methods for delivery of evidence-based psychotherapies: A systematic review.Implementation Science. doi: 10.1186/s13012-020-00998-w.

[55]. Ma C., Huang J. B.,Yang X., and Yang M. H. (2018). Adaptive Correlation Filters with Long-Term and Short-Term Memory for Object Tracking.Int J Comput Vis.doi: 10.1007/s11263-018-1076-4.

[56]. Chen Y. (2021). Evaluation of teaching effect of internet of things education platform based on long-term and short-term memory network.Int J Contin Eng Educ Life Long Learn.doi: 10.1504/IJCEELL.2021.111839.

[57]. AlpagoD., ZorziM., and Ferrante A. (2020). Link Prediction: A Graphical Model Approach. doi: 10.23919/ecc51009.2020.9143706.

[58]. Perrot N. et al.(2015). A decision support system coupling fuzzy logic and probabilistic graphical approaches for the agri-food industry: Prediction of grape berry maturity.PLoS One.doi: 10.1371/journal.pone.0134373.