



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 5, May 2021

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 7.512**

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)

# Intelligent Sales Prediction using Machine Learning

Taman Poojary<sup>1</sup>, Deepak Ramchandani<sup>2</sup>, Ismail Pawaskar<sup>3</sup>, Nikita Lemos<sup>4</sup>

U.G. Student, Department of Information Technology, Xavier Institute of Engineering, Mahim, Maharashtra, India<sup>1</sup>

U.G. Student, Department of Information Technology, Xavier Institute of Engineering, Mahim, Maharashtra, India<sup>2</sup>

U.G. Student, Department of Information Technology, Xavier Institute of Engineering, Mahim, Maharashtra, India<sup>3</sup>

Professor, Department of Information Technology, Xavier Institute of Engineering, Mahim, Maharashtra, India<sup>4</sup>

**ABSTRACT:** Intelligent Decision Analytical System requires combination of decision exploration and prediction. The greater part of the business associations intensely rely upon an information base also, request expectation of deals patterns. The precision in deals figure gives a major effect in business. Pattern Analytical System and Data Mining techniques are powerful tools that helps in removing covered up information from a huge dataset to upgrade exactness and productivity of estimating. Traditional prediction frameworks are hard to manage the enormous information and exactness of sales measure. Valuable metals like diamonds and gold are sought after because of their financial prizes. In this manner, different methods are commonly utilized to estimate costs of jewels and valuable metals with the point of quick and precise outcomes. The costs vacillate day by day making it hard to predict the next future worth. Subsequently, by analyzing the example of past costs we can apply relapse models for future expectation. This paper targets forecasting the future costs of valuable metals like diamond, utilizing various machine learning techniques, intending to get the most precise consequence of all. Group models are utilized for expanding the exactness of costs.

**KEYWORDS:** Machine Learning, Prediction, Data Mining, Analytics, Estimation, Pattern.

## I. INTRODUCTION

One of the critical objections of this assessment work is to find the trustworthy arrangements design gauge instrument which is executed by using information mining examination strategies to achieve the best pay. The current business handles immense storage facility of data. The volume of data is depended upon to turn out to be further in a dramatic manner. The actions are necessary in order to oblige measure speed of trade and to update the typical advancement in data volume and customer lead. For an impressive period of time, human headways from one side of the planet to the other have wanted valuable stones for their elegant brilliance and enchanting shimmer [1].

Seen as a picture for wealth, they are assessed for the wealthy by the jewel adventures that overpower the market. While these valuable stone producers and promoters have been incredibly compelling previously, with the start of the information age, buyers are as of now proficient to cross-reference expenses of practically identical valuable stones from different associations prior to making purchasing decisions. Lab evaluated quality can be speedily settled for jewels additionally, promising the request for the attributes of a valuable stone [2]. With these assets comes the assignment of assessing the cost of cut precious stones given their properties. There are many factors that might affect the price of a diamond, but the most common ones are referred to as the 4 Cs: carat, cut, color, and clarity.

**Carat:** Carat is the mass of the diamond. 1 carat (ct) is equivalent to 200mg. This is the main quantitative proportion of the 4 Cs. Carat is non-straight identified with the cost.

**Cut:** Cut alludes to both the state of the stone and the nature of its glitter. The cut flawlessness is grouped from "Fair" to "Ideal".

**Color:** Precious stone tints vary from dismal to a light yellow. The more insipid a valuable stone is, the more expensive it is most likely going to be. The standard is a portrayal made by the Gemmological Establishment of America and is



the most used out of the whole of the concealing auditing plans; it uses an in consecutive request score, "D" being the most vapid and "Z" being a prominent yellow.

**Clarity:** Diamonds may have inside imperfections and breaks which decline their straightforwardness, which thus diminishes their worth. Lucidity is evaluated on a scale from FL (Flawless) to I3 (Obvious Inclusions) in light of the size, nature, position, and amount of inward imperfections.

There are a couple of various properties that may impact the expense of a valuable stone. While these likely will not be also known as the customarily used 4 Cs, they in spite of everything may measure a couple of factors of the valuable stone that could impact the quality and hence the expense of the gems.

**Depth:** Depth is estimated as the proportion among z and the normal width of the head of the diamond.

**Table:** Table is estimated as the width of the head of the diamond at its most stretched out point. The dataset contains data on costs of jewels, just as different credits of precious stones, some of which are known to impact their cost: the 4 Cs (carat, cut, color, and clarity) , as well as some physical measurements (depth, table, price, x, y, and z) [6]. The figure below shows what these measurements represent.

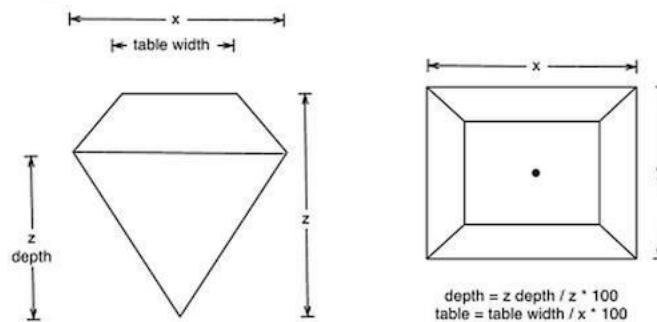


Fig.1: Exploring dimensions of Diamond

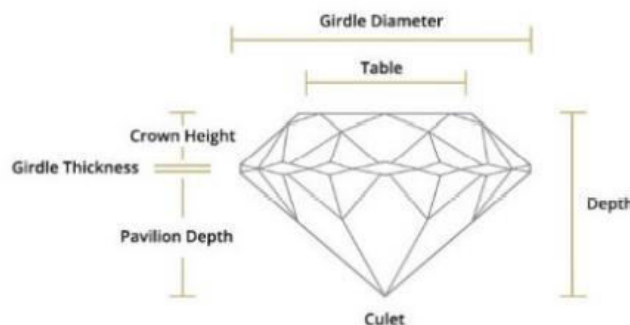


Fig.2: Structure of the diamond

The carat is estimated in carats, the cut is estimated subjectively as an evaluation from "Fair", "Good", "Very Good", "Premium", and "Ideal", the shading is estimated with grades from "J" being the most exceedingly awful to "D" being the best, the clarity is estimated utilizing the standard from "I1", "SI2", "SI1", "VS2", "VS1", "VVS2", "VVS1", and "IF", the depth is measured as the proportion of the diamond's z axis to its normal distance across, the table is estimated as the width of the diamond at its broadest point, the directions are estimated in millimetres, and the price is estimated in US dollars [7].

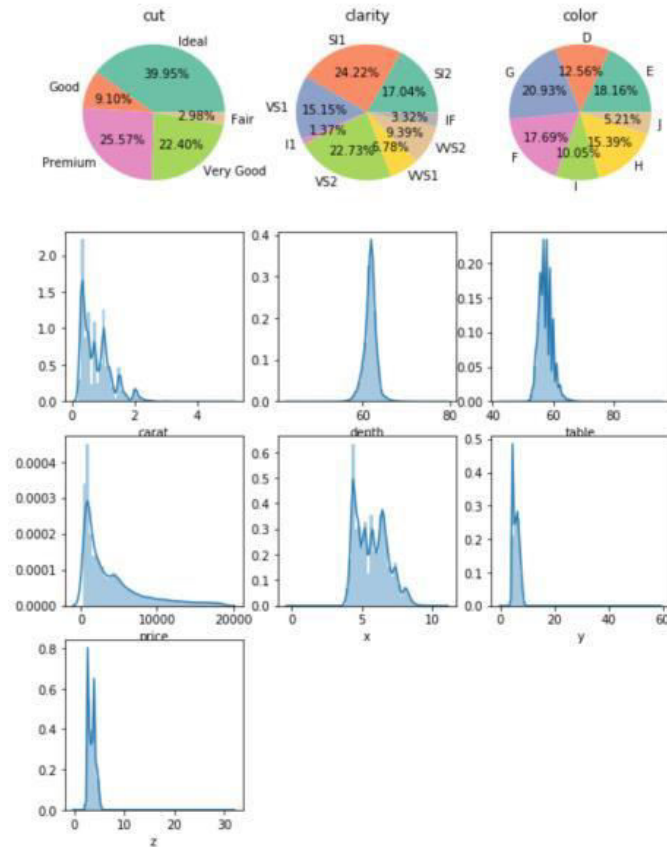


Fig.3: Distributions of data on different properties

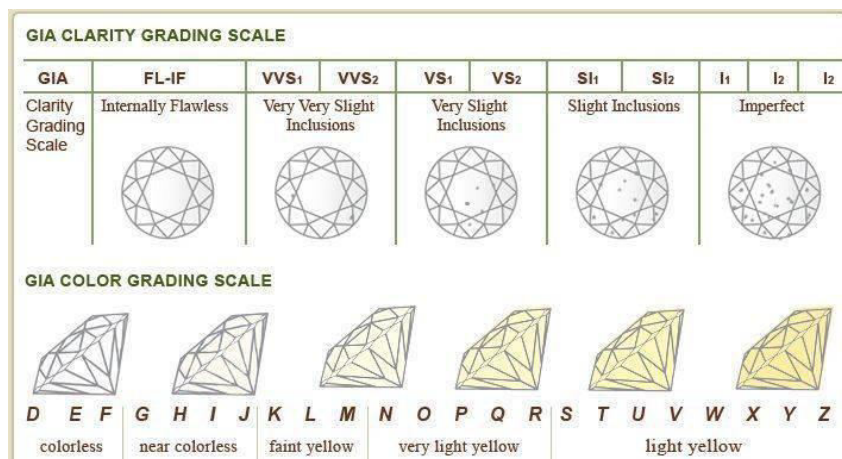


Fig.4: Diamond Chart based on color and clarity

## II. RELATED WORK

### Literature Survey:

Accurate estimates grant the relationship to improve promote advancement with more raised degree of pay age. Data mining systems are especially incredible in tuning gigantic volume of data into supportive information for cost assumption and arrangements gauge, it is the key of sound arranging. Special case area this methodology plays out each major datum pre-taking care of and model smoothing out. Exemption acknowledgment system can be used to



pass on the model or as a starting point for extra improvements and steady in showing customary information which is liberated from the models. The oversupply of harsh stones and the undeniably stressed funds of mediators have hit excavators' monetary records as of late as they attempt to deal with the excess and increment the benefit of existing stones [1].

Gradient Boosted Trees Gradient boosting is a machine learning technique for regression and classification problem. This approach could ensemble learning method that combines large number of decision trees to produce final prediction model. This model is built on a principle that a collection of weak learners combined together can produce a strong learner by using boosting process. Traditional estimating strategies face difficulties in delivering precise deals information for new items and buyer situated merchandise. [2].

The data input squares are shown by the "Recover" head. The "Recover" chairman stacks a Quick Excavator object into the data stream measure. In the specific case it license to pick data set aside in the close by store isolated by the csv preliminary dataset. In order to structure a commonsense planning dataset, a reasoning affiliation must applied on every one of the three datasets, procuring a single dataset to measure [3].

This investigation paper is as for the business conjecture reliant on current clothing style. In current society being stylish is an example. There are various components sway clothing and style bargains, including esteem, sort of material, and size, alongside the customary factors like season and material. This investigation paper used decision tree and ID3 computation to lead the assumption assessment regarding the dress arrangements [4].

Solid assumptions benefit by having superior grade and successfully open data. Dependent upon the aftereffect of the association, different sorts of outside data could be used. Thusly, it is critical not only to make figures reliant on the numbers nearby yet notwithstanding match these numbers with emotional information in order to get an undeniably functional viewpoint on the business . This can be refined with fitting correspondence and collaboration between the businessand the gathering related with the improvement of the assessing model [5].

In this paper, it is realized a structure to evaluate the idea of framework resources, which accepts four man-made intelligence computations to calculate the pack scores. The whole model contains various portions: data preprocessing, model getting ready and assumption using the estimations like Arbitrary Timberland KNN and Calculated Relapse. We took a gander at the introduction of four assumption systems on a dataset with 100 ordinary programming packs. It shows that these techniques have a high exactness, which is from 82.84% to 90.52%. [6].

In this paper, we propose a period revelation computation composed example estimate. The assumption capacity is refined by learning automata. Another example assumption and period revelation computation is proposed in this paper. To the furthest extent that we might actually know, this is the principle estimation that can normally find the hour of data, which is furthermore used to expect its future example [7].

In order to adjust to the multifaceted design of hardware plan smoothing out, we address the issue by starting from the assessment of one of the target features that is the domain of a gear portion. To achieve this level headed, we resort to different AI computations to play out the assumption by using estimations like Arbitrary Woods and Inclination Lift. Three ML computations have been considered to predict the zone of a RI part [8].

Especially in clinical field, where those methodologies are comprehensively used in assurance and assessment to settle on decisions by using Backing Vector Machines, Innocent Bayes and KNN. In this examination, we used four basic estimations: SVM, NB, k-NN and C4.5 on the Wisconsin Bosom Disease (remarkable) datasets [9].

With the headway of huge data assessment advancement, more thought has been paid to disorder gauge according to the perspective of huge data examination, diverse investigates have been driven by picking the characteristics normally from a gigantic number of data to improve the accuracy of risk gathering, rather than the as of late picked characteristics [10].

Consequently, we can arrive at the derivation that the exactness of CNN-UDRP (T-data) and CNNMDRP (S&T-data) computations have little qualification anyway the audit of CNN-MDRP (S&Tdata) estimation is higher and its



association speed is speedier. In summation, the introduction of CNN-MDRP (S&T-data) is superior to CNNUDRP (T-data) [10].

AI techniques can be applied to all orders. Computer based intelligence uses experiences to clarify various portrayal and packing issues. The ML computations are described in three orders . They are overseen, independent and semi coordinated. In this paper we analyzed around three simulated intelligence computations which can be applied to gauge, as Generalized Linear Model (GLM), Decision Tree (DT) and Gradient Boost Tree (GBT) [11].

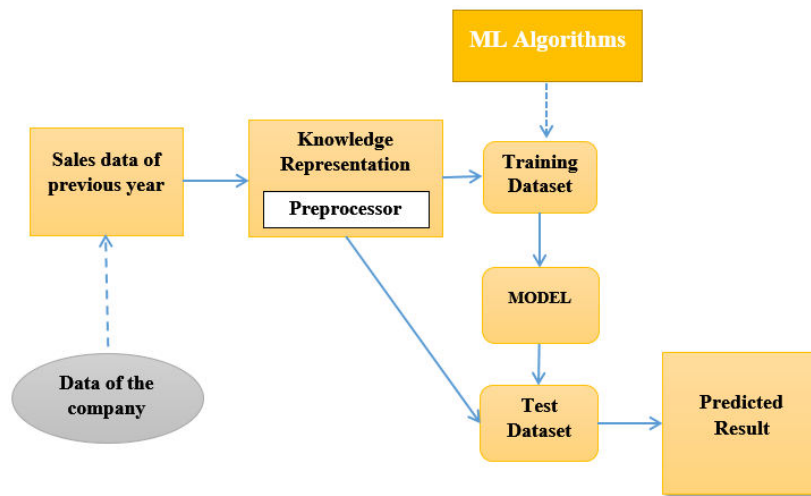


Fig.5: System Architecture

To gauge the electric usage, chronicled data on the past use, temperature, wetness, absolute public yield (Gross domestic product), people, sorts of families and their association coefficients, and customer rehearses are routinely utilized. The lead projection is known to use the data of people's inclinations in using electrical contraptions (which isn't associated with the biological conditions and the environment) for assumption. [12].

### III. METHODOLOGY

#### DATA PREPROCESSING:

This model is utilized to pre-measure the information like eliminating the repetitive credits or records, information purging, sifting the dataset, and so on Mathematical factors can be delegated persistent or discrete dependent on whether the variable can take on a limitless number of qualities or just non-negative entire numbers, individually. On the off chance that the variable is all out, we can decide whether it is ordinal dependent on whether the levels have a characteristic requesting [11].

Information Purifying or data cleaning is the path toward perceiving and changing (or removing) degenerate or wrong records from a record set, table, or data set and insinuates recognizing divided, mixed up, misguided or pointless bits of the data and a while later overriding, changing, or deleting the muddled or coarse data. Data cleansing may be performed brilliantly with data battling gadgets, or as gathering getting ready through scripting [12].



carat	0
cut	0
color	0
clarity	0
depth	0
table	0
price	0
x	0
y	0
z	0

As it tends to be seen here, none of the attributes has null qualities. Hence, interaction of topping off the invalid qualities or computing intend to fill it very well may be overlooked.

**EXPLORATORY DATA ANALYSIS:**

Authentic Data Examination by and large fuses a type of quantifiable gadgets, which a layman can't perform without having any quantifiable information. There are unmistakable programming packs to perform obvious information appraisal. This thing wires Quantifiable Assessment Structure , Genuine Group for the Humanistic systems, Detail delicate, and so on. After information pre-processing, to plainly see the value in the chance of our information, an exploratory assessment was coordinated [13]. Abnormality Area strategy plays out every essential datum pre-processing moreover, model movement.

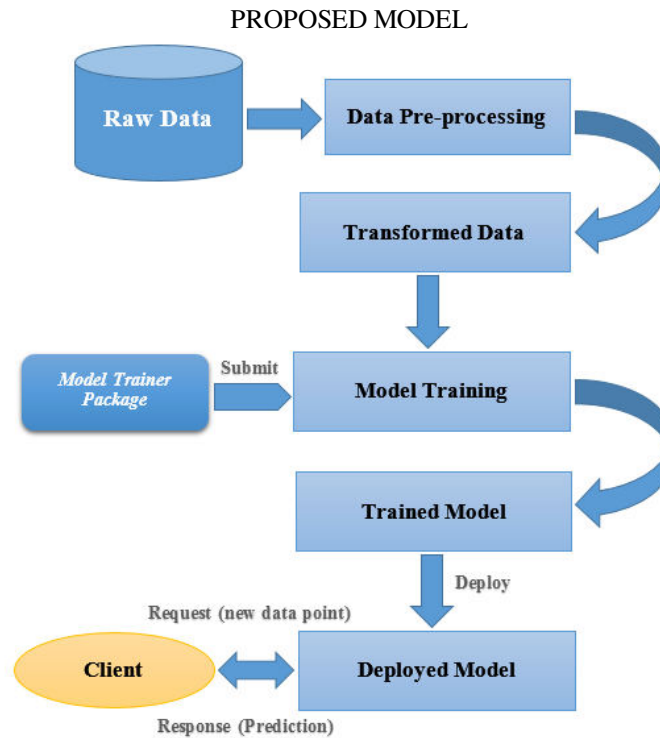


Fig.6: System Flow

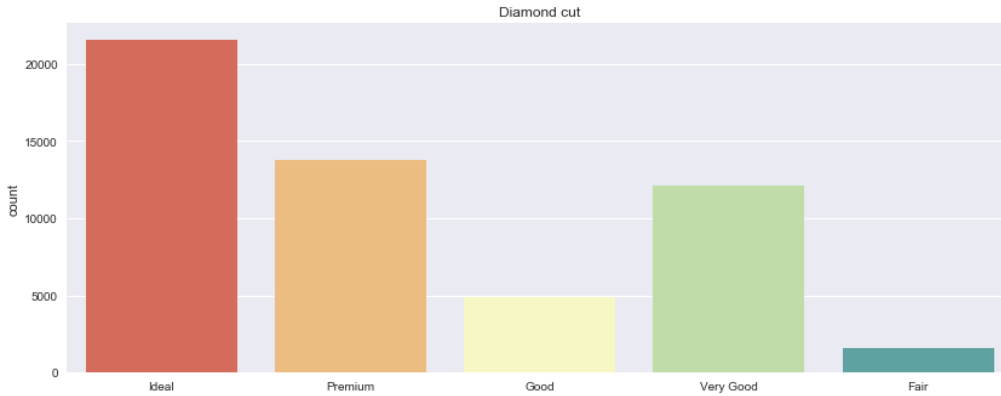


Fig.7: Count of diamond based on cut

Here, it can be seen that count of the ideal diamond is more. Our goal was to anticipate the expense of precious stones using features, for instance, carat, clearness, shading, cut, profundity, table length, x, y, and z pivot lengths in millimeters. The profundity, table length, x, y, and z rotate lengths were given as mathematical data. Thusly, we normalized the data by removing the mean from each data point and disengaging by the standard deviation.

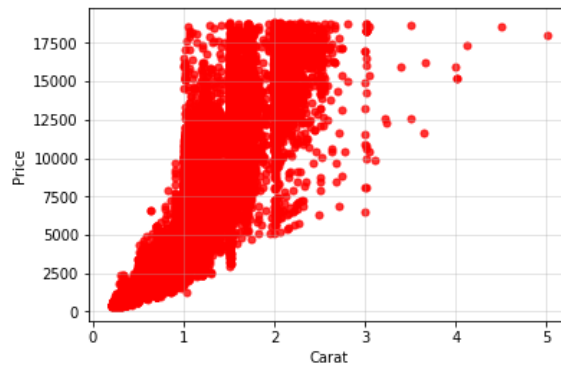


Fig.8: Scatterplot of Carat vs Price

As the carat of the diamond continues to expand thus as the cost. AI strategies can be applied to all controls. Computer based intelligence uses experiences to settle various portrayal and bundling issues. The ML estimations are described in three characterizations. They are managed, independent and semi directed. In this paper we inspected around three computer based intelligence estimations which can be applied to gauge.

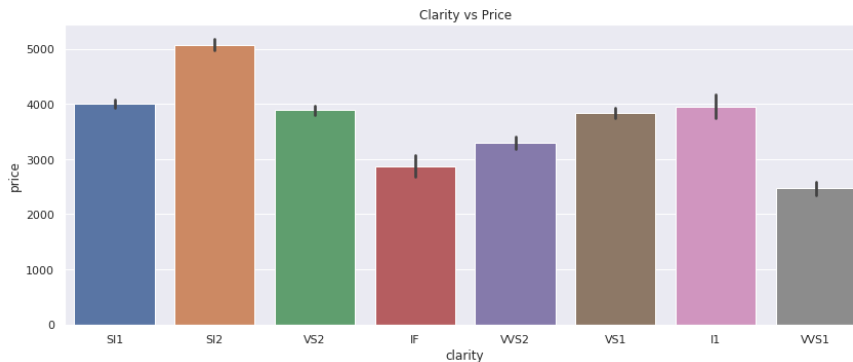


Fig.9: Histogram of clarity vs price



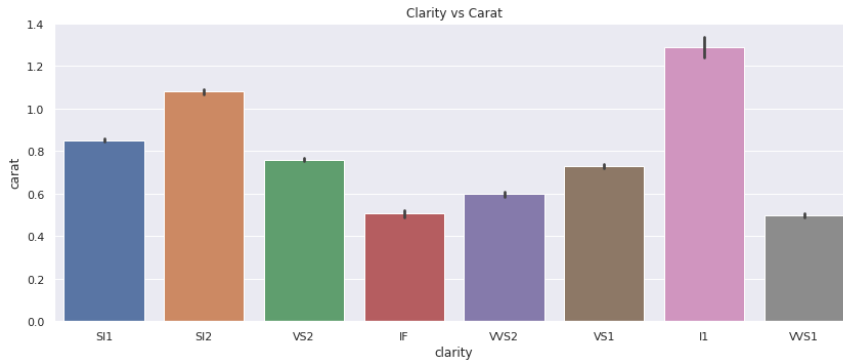


Fig.10: Histogram of clarity vs carat

Here, it can be seen that J color diamonds are the most expensive and heaviest diamonds. The two plots are very similar. Thus, it can be concluded that the heavier diamond is expensive, if only color is considered. The carat, clarity, color, what's more, cut were given as outright data so we changed over it to a component vector by arranging them to a continuous number run with zero mean. For example, the cut component of a precious stone, is addressed as Reasonable, Great, Generally excellent, Premium, or then again Ideal where Reasonable is the most really dreadful while Ideal is the awesome.

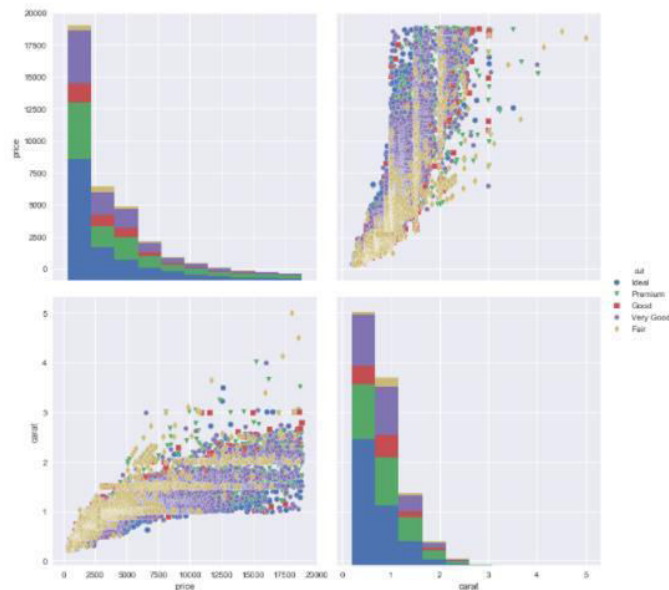


Fig.11: Comparison of carat with price based on diamond cut

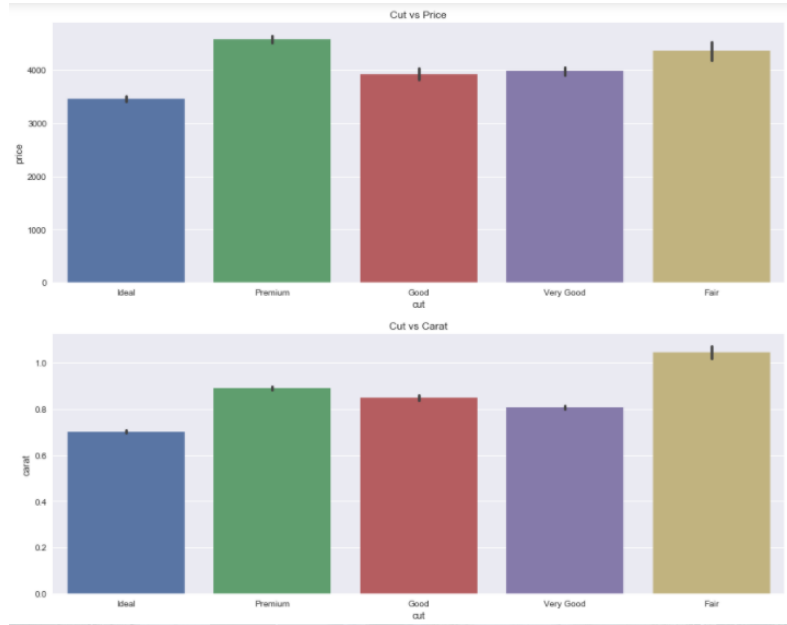


Fig.12: Histograms based on the diamond cut

From here, the conclusion can be drawn as

- ⊣ Fair cut diamonds weigh the most but are not the most expensive diamonds.
- ⊣ Premium cut diamonds are the most expensive diamonds.
- ⊣ Ideal cut diamonds weigh less and are cheapest diamonds.

We can see that price of diamond is dependent on the cut.

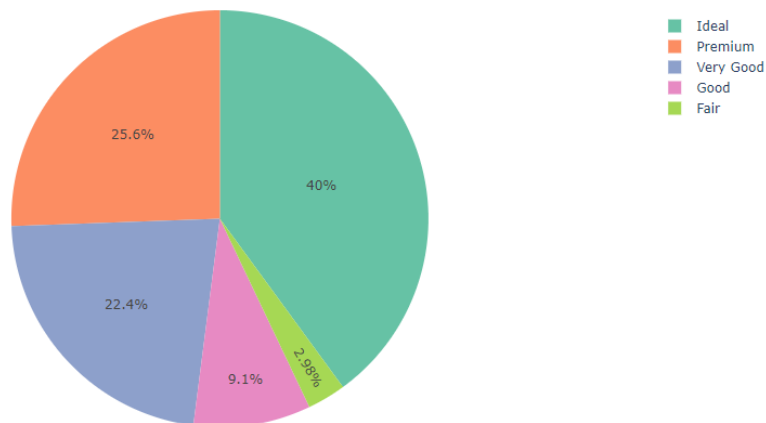


Fig.13: Pie Chart for count of Diamonds

From the above figure, it is clearly seen that the mean the Ideal jewels is 40% which is the most elevated of all. Fair Cut precious stones which weigh the most are smallest in number which is almost around 3%.

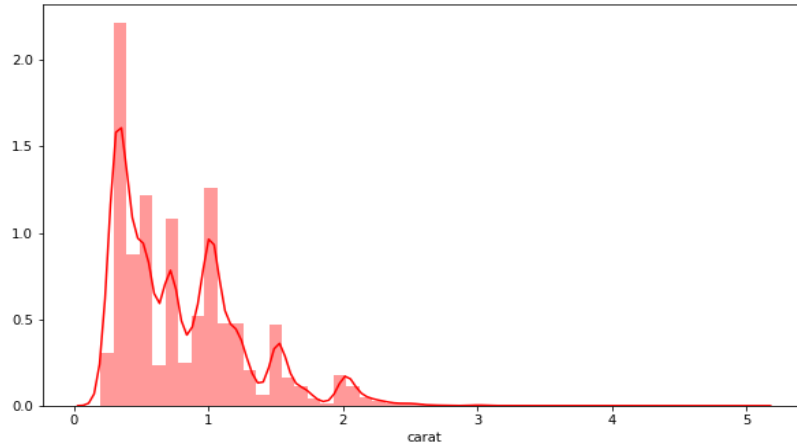


Fig.14: Plot for the count of diamonds based on carat

Here, it is seen that the count of the diamonds of the carat between 0 and 1 is the highest

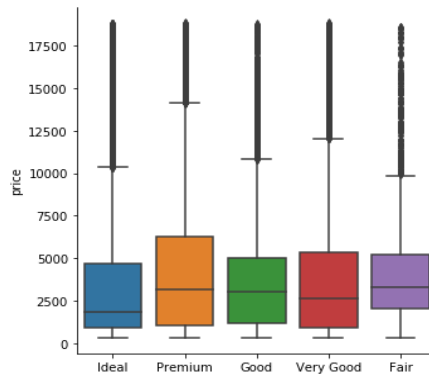


Fig.15: Boxplot of cut vs price of diamond

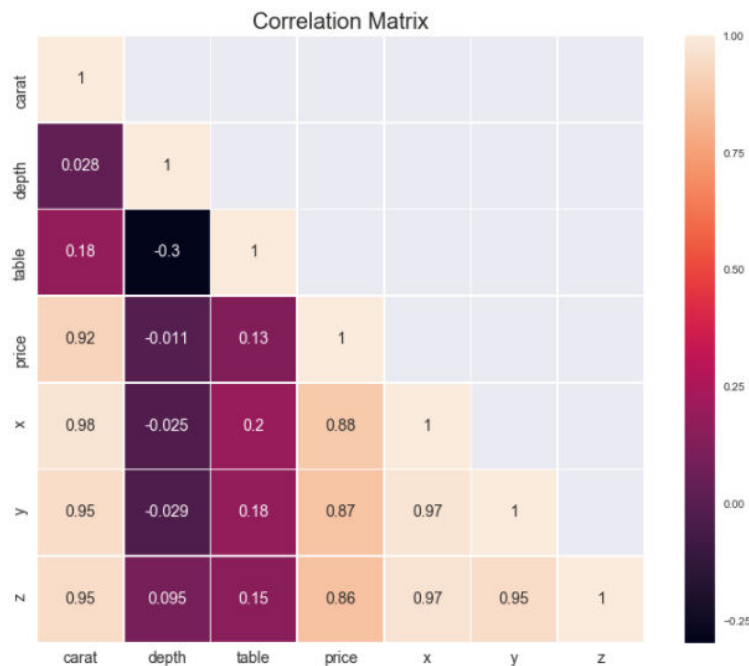


Fig.16: Correlation Matrix between features of diamond



The conclusion can be drawn are as follows:

- L Depth and Table have almost no correlation with price of diamond and therefore it can be dropped.
- L Price of the diamond and carat weight are highly correlated.
- L Carat is one of the main features to predict price of diamond.
- L There seems to be some multi-collinearity with some of the independent variables namely carat, x, y and z.

It's worthwhile to recognize that cost is firmly directly corresponded with the carat of a precious stone, while not impeccably associated. We realize since we can utilize carat as a solid indicator at the cost of a diamond. The heatmap doesn't uncover everything about the information, notwithstanding. Other than carat, the size, and the cost of the jewel, there doesn't appear to be any direct connection, yet, on the off chance that we plot the information outwardly while naming the downright properties, we can see there is a type of example

	carat	cut	color	clarity	price	x	y	z
0	0.23	4	5	1	326	3.95	3.98	2.43
1	0.21	3	5	2	326	3.89	3.84	2.31
2	0.23	1	5	4	327	4.05	4.07	2.31
3	0.29	3	1	3	334	4.20	4.23	2.63
4	0.31	1	0	1	335	4.34	4.35	2.75

Here some multicollinearity is suspected. It can be confirmed by checking the Variance Inflation Factor (VIF). Each independent variable is regressed against each independent variable and the VIF is calculated.

Variance Inflation Factor (VIF) is a proportion of the measure of multicollinearity in a bunch of different relapsefactors. Numerically, the VIF for a relapse model variable is equivalent to the proportion of the general model difference to the change of a model that incorporates just that solitary free factor.

$$VIF = \frac{1}{1-R^2}$$

If the independent variable is not dependent on other independent variables, an acceptable VIF will fall between 1 and 5. A VIF greater than 10 is a clear indication that there is some collinearity between the variables.

Features	VIF
Intercept	68.412939
carat	1.057123
cut	1.102490
color	1.215142
clarity	36.063353
x	20.491380
y	21.269593
z	

Such high VIF scores confirm the suspicion! The variables x, y and z are dropped leaving only price, carat, cut, color and clarity.

**RESERCH METHODOLOGIES:**

The principle motivation behind this exploration is to assess and investigate the utilization of information digging strategies for deals gauging, to deliver models which are far reaching and dependable.

**1. Linear Regression**

It is helpful for discovering connection between two constant factors. One is indicator or autonomous variable and other is reaction or ward variable. It searches for measurable relationship yet not deterministic relationship. Connection



between two factors is supposed to be deterministic on the off chance that one variable can be precisely communicated by the other. For instance, utilizing temperature in degree Celsius it is feasible to precisely anticipate Fahrenheit. Factual relationship isn't exact in deciding connection between two factors.

$$y = \alpha + \beta x$$

$\beta$  = slope

$\alpha$  = y-intercept

$y$  = y- coordinate

$x$  = x-coordinate

### 2. Polynomial Regression

In statistics, polynomial regression is a form of regression analysis in which the relationship between the independent variable  $x$  and the dependent variable  $y$  is modelled as an  $n$ th degree polynomial in  $x$ . Polynomial regression is a kind of relapse wherein the association between the a free factor  $x$  and the reliant variable  $y$  is shown as a farthest breaking point polynomial in  $x$ . Polynomial backslide fits a nonlinear association between the assessment of  $x$  and the contrasting unexpected mean of  $y$ , shown  $E(y | x)$ . But polynomial backslide fits a nonlinear model to the data, as an authentic assessment issue it is straight, as in the backslide work  $E(y | x)$  is immediate in the dark limits that are evaluated from the data.

$$Y = \beta_0 + \beta_1 X_j + \beta_2 X_j^2 + \epsilon$$

$$Y = \beta_0 + \beta_1 X_j + \beta_2 X_j^2 + \dots + \beta_p X_j^p + \epsilon$$

where:

$\beta_0$  = Population regression constant

$\beta_j$  = Population regression coefficient for variable  $X_j : j = 1, 2, \dots, k$

$p$  = Order of the polynomial

$\epsilon_j$  = Model error

### 3. Decision Tree Regressor

Decision Tree algorithm has become one of the most used machine learning algorithm both in competitions like Kaggle as well as in business environment. Decision Tree can be used both in classification and regression problem. This article present the Decision Tree Regression Algorithm along with some advanced topics. The decision of making strategic splits heavily affects a tree's accuracy. The decision criteria is different for classification and regression trees. Decision trees regression normally use mean squared error (MSE) to decide to split a node in two or more sub-nodes.

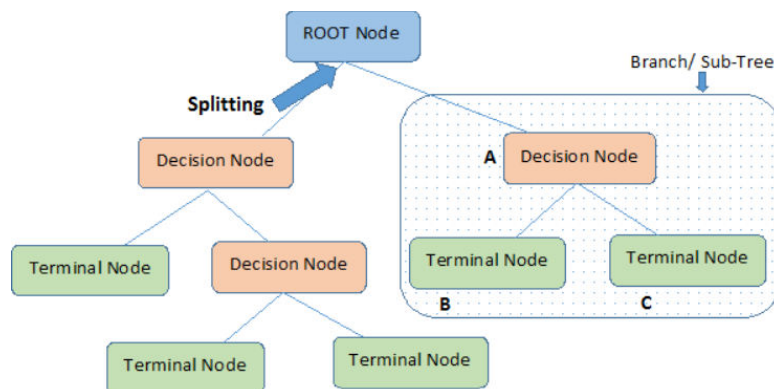


Fig.17: Structure of Decision Tree



#### 4. Random Forest Regressor

A Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as bagging. The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees. Random Forest has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model. This part is called Bootstrap.

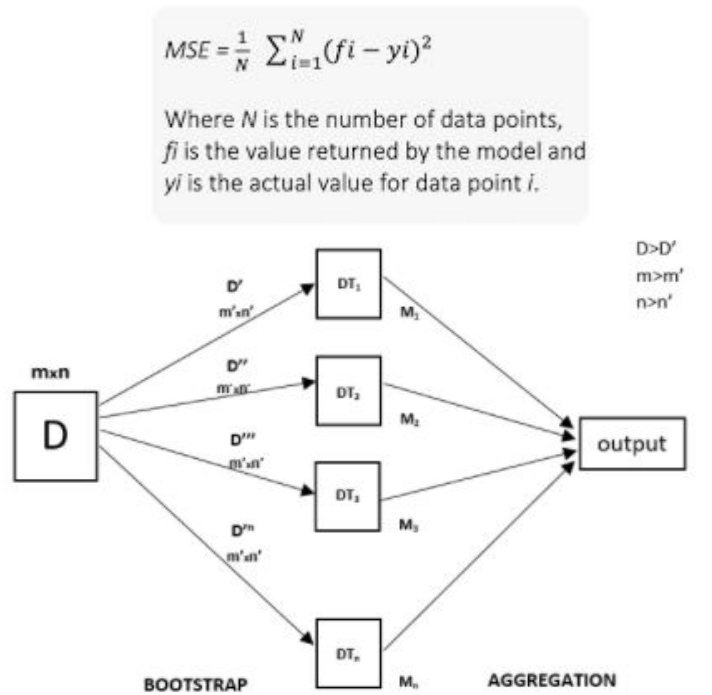


Fig.18: Structure of Random Forest Regressor

#### 5. Gradient Boost Regressor

Gradient Boosting is a machine learning procedure for regression and characterization issue. This methodology could be a learning technique that joins huge number of choice trees to create last forecast model. This model is based on a rule that an assortment of powerless students consolidated together can create a solid student by utilizing boosting measure. It delivers an expectation model as a gathering of powerless forecast models, ordinarily choice trees. It assembles the model in a phase savvy design like other boosting strategies do, and it sums them up by permitting improvement of a discretionary differentiable misfortune work.

$$y_i^p = y_i^p + \alpha * \delta \sum (y_i - y_i^p)^2 / \delta y_i^p$$

which becomes,  $y_i^p = y_i^p - \alpha * 2 * \sum (y_i - y_i^p)$

where,  $\alpha$  is learning rate and  $\sum (y_i - y_i^p)$  is sum of residuals

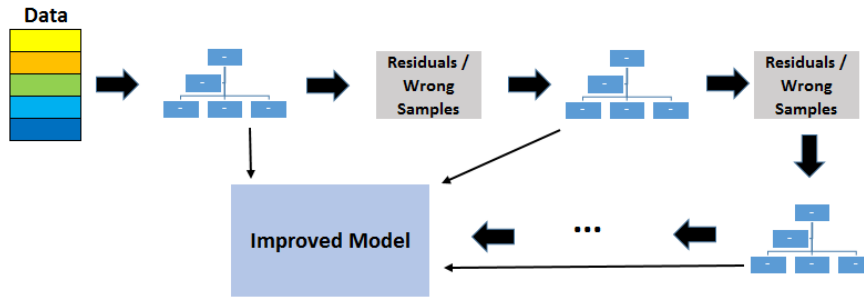


Fig.19: Gradient Boost Model

#### IV. EXPERIMENTAL RESULTS

The performance of the algorithms is mainly focused on accuracy. Accuracy in each class and confusion matrix which shows the number of predictions of each class which can be compared to the instances of each class. Root Mean Square Error, Mean Square Error, Absolute error are calculated and average of the error.

On a typical the immediate backslide scored the exactness around 90% by eliminating the fundamental credits when required.

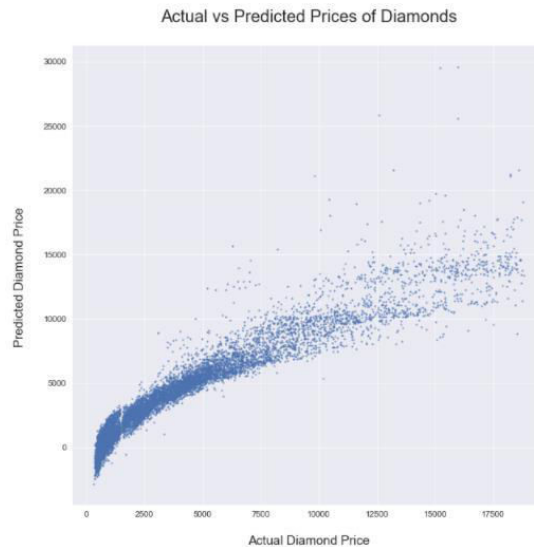


Fig.20: Scatterplot based on Linear Regression

Later, we determine the model and improve its performance by plotting it and finding the degree at which it gives the best results and the degree was found to be 5 with the highest R2 value of 0.97.

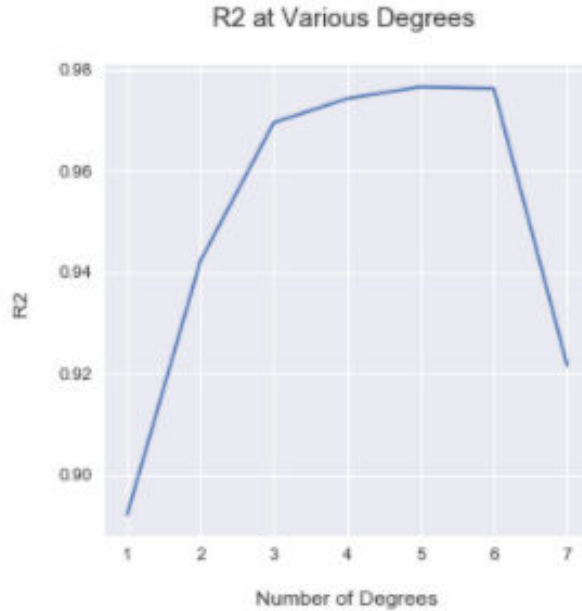


Fig.21: R2 value at various degrees

The model is prepared on 43,152 instances (80%) and tried on the leftover 10,788 occurrences (20%). A limited quantity of L2 penalty was applied to diminish a portion of the fluctuation presented by a serious level polynomial calculation. K-fold cross validation with 4 folds gives a steady presentation. This model can almost certainly sum up well to new information. The R2 value approaches to 97% with a root mean squared error of 610, approximately 15% of the mean of the target variable.

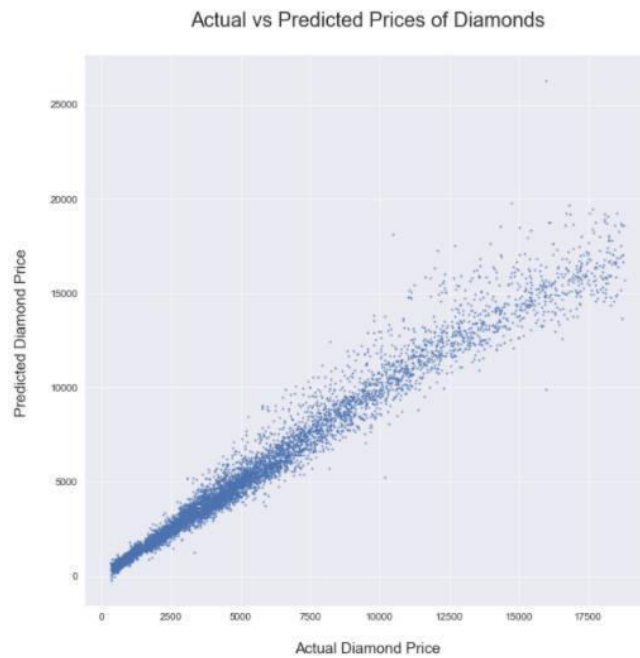


Fig.22: Scatterplot based on Polynomial Regression

As it can be seen in the above figure the model is very stable compare to the previous one. It is easy to see that this model has much preferred prescient abilities over the past one. Whereas there are some outliers as the model goes linearly upwards.

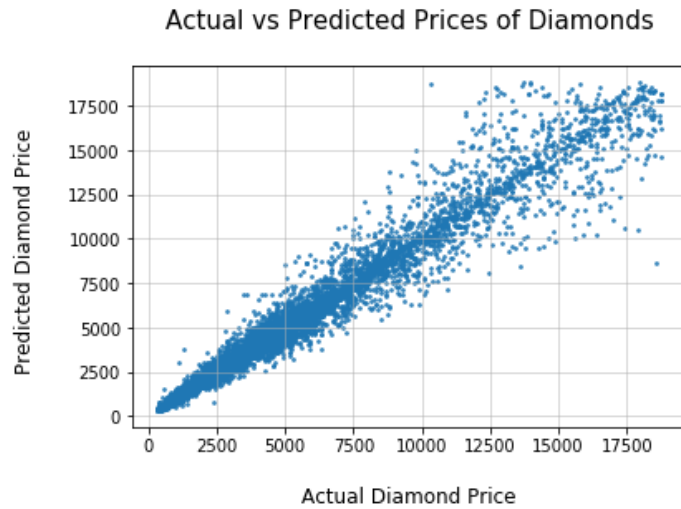


Fig.23: Scatterplot based on Decision Tree Regressor

Here, the model gives the accuracy of around 96.5 and it contains some outliers as the model proceed forwards.

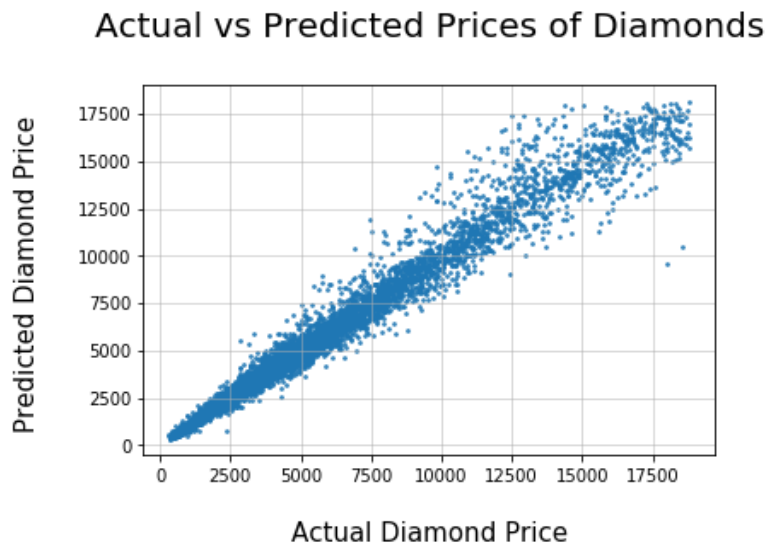


Fig.24: Scatterplot of Random Forest Regressor

This model accomplishes the exactness of 97.92% which is considerably more steady than different ones with the cross validation of  $k=4$ . Additionally the Gradient Boost Regressor achieves the exactness of 97.95% which is a lot of firmly identified with that of Random Forest Regressor.



Table I: Performance Summary

Model Name	Accuracy Rate (%)
Linear Regression	90.06
Polynomial Regression	97.64
Decision Tree Regressor	96.54
Random Forest Regressor	97.92
Gradient Boost Regressor	97.94

From the above table, it is seen that Random Forest and Gradient Boost gives the higher accuracy which suites the dataset most for prediction.

## V. CONCLUSION

The chronicled scenery of man-made thinking gives us that there has been a lethargic and trans-developmental improvement inside the uncommon pieces of computational sciences that underlie the advances regular in computerbased intelligence. A huge bit of the advancements involve procedures portrayed by so-called computational understanding, including neural frameworks, formative estimations and cushioned systems. Regardless of the way that it may have all the earmarks of being unavoidable that such a historic business instrument will be embraced as once tremendous crowd, in reality more nuanced than that. The researchers have concluded that an intelligent sales prediction system is required for business organizations to handle enormous volume of data. Business decisions are based on speed and accuracy of data processing techniques. Machine learning approaches highlighted in this research paper will be able to provide an effective mechanism in data tuning and decision making. In order to be competent in business, organizations are required to equip with modern approaches to accommodate different types of customer behavior by forecasting attractive sales turn over.

Computer based intelligence model assumptions license associations to make significantly precise guesses in regards to the likely aftereffects of a request reliant on bonafide data, which can be basically a wide scope of things – customer disturb likelihood. As an issue of first significance, it is a benchmark. We can use it as old news level we will achieve if nothing changes in our technique. Plus, we can register the consistent assessment of our new exercises on top of this benchmark. It will in general be utilized for masterminding. We can plan our advantage and nimbly exercises by looking at the checks. It helps with seeing where to offer more.

## REFERENCES

- [1] Sales Prediction using Machine Learning, Sunitha Cheriyan, Shaniba Ibrahim, Saju Mohanan, Published 2018 Computer Science 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE)
- [2] A survey on retail sales forecasting and prediction in fashion markets, Systems Science & Control Engineering: An Open Access Journal, 3:1, 154-161, DOI: 10.1080/21642583.2014.999389
- [3] Data Mining Model Performance of Sales Predictive Algorithms Based on Rapidminer Workflows, Alessandro Massaro, Vincenzo Maritati, Angelo Galiano Published 2018 Computer Science International Journal of Computer Science and Information Technology
- [4] The Implementation of Data Mining Techniques for Sales Analysis using Daily Sales Data in International Journal of Advanced Trends in Computer Science and Engineering 8(1.5):74-80 · November 2019 DOI: 10.30534/ijatcse/2019/1681.52019





- [5] Deloitte Sales Forecasting Deloitte Analytics Approach The growing world of data [https://www2.deloitte.com/content/dam/Deloitte/it/Documents/technology/Sales%20forecasting\\_Deloitte%20Analytics%20Approach\\_Deloitte%20Italy.pdf](https://www2.deloitte.com/content/dam/Deloitte/it/Documents/technology/Sales%20forecasting_Deloitte%20Analytics%20Approach_Deloitte%20Italy.pdf)
- [6] Resource Quality Prediction Based on Machine Learning Algorithms 4th International Conference on Systems and Informatics (ICSAI 2017) DOI: 110.1109/Cybermatics\_2018.00161 2017 4th International Conference on Systems and Informatics (ICSAI)
- [7] Period Detection and Future Trend Prediction Using Machine Learning Techniques 21st Euromicro Conference on Digital Systems and Electronics Conference: 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom)
- [8] A Machine Learning Approach for Area Prediction of Hardware Designs from Abstract Specifications Volume 71, November 2019, 102853 published at IEEE Machine Learning Design productivity Area estimation
- [9] Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis Volume 83, 2016, Pages 1064-1069 published at Elsevier Procedia Computer Science DOI: 10.1109/CIMCA.2018.8739696 published at IEEE.
- [10] Disease Prediction by Machine Learning over Big Data from Healthcare Communities International Journal of Innovative Research in Computer and Communication Engineering Vol. 5, Issue 12 December 2017 DOI: 10.1109/ACCESS.2017.2694446
- [11] Intelligent Sales Prediction using Machine Learning Techniques 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, August 2018.
- [12] A Machine Learning-based Approach for The Prediction of Electricity Consumption Volume 134, December 2019 published at Elsevier State-of-the-art approaches to estimate energy consumption in machine learning
- [13] A PSS model for diamond gemstone processing: economic feasibility analysis /Forty Sixth CIRP Conference on Manufacturing Systems 2017 published by Elsevier B.V 300A, box 2422, 3001 Leuven, Belgium doi: 10.1016/j.procir.2013.06.005.



**Impact Factor:  
7.512**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details