



# International Journal of Engineering Research and Science & Technology

[www.ijerst.org](http://www.ijerst.org)

ISSN : 2319-5991

Vol. 21 No. 3 (1) 2025



[ijerst.editor@gmail.com](mailto:ijerst.editor@gmail.com)  
[editor@ijerst.com](mailto:editor@ijerst.com)

**Research Paper****HYBRID DIAMOND CUT CLASSIFICATION AND PRICE PREDICTION USING FFNN BASED RF-CART FRAMEWORK**Pranantha Katkuri<sup>1</sup>, Talla Sushma<sup>2</sup>, Karamareddy Sharmila<sup>3</sup><sup>1</sup>M. Tech Student, <sup>2</sup>Assistant Professor, <sup>3</sup>Associate Professor & Head, <sup>1,2,3</sup> Department of Computer Science and Engineering<sup>1,2,3</sup> Vaagdevi Engineering College, Bollikunta, Warangal, 506005, TelanganaEmail - <sup>1</sup>[pranantha5@gmail.com](mailto:pranantha5@gmail.com), <sup>2</sup>[sushmavisu410@gmail.com](mailto:sushmavisu410@gmail.com), <sup>3</sup>[sharmilakreddy@gmail.com](mailto:sharmilakreddy@gmail.com)**ABSTRACT**

Diamonds are a highly valued commodity, where accurate quality classification and price prediction are critical for both consumers and retailers. Market analytics indicate that over 80% of diamond transactions occur via digital platforms, yet more than 40% of price estimations lack consistency due to subjective grading and market volatility. Existing systems often struggle to map categorical quality indicators, such as *cut*, to price, hindered by class imbalance and non-linear data complexities. The dataset used in this study includes attributes such as carat, cut, color, clarity, depth, table, x, y, z, and price, with *cut* as the classification target. Traditional models show limited performance due to skewed class distributions and weak feature interactions. To address these challenges, an advanced pipeline was developed. Exploratory Data Analysis (EDA) was conducted to uncover patterns and correlations between quality attributes and price. Data preprocessing ensured normalization and encoding of categorical features, while Support Vector Machine–Synthetic Minority Oversampling Technique (SVM-SMOTE) was applied to balance the five cut classes: Fair, Good, Ideal, Premium, and Very Good. For predictive modelling, several hybrid Classification and Regression Tree (CART) models were tested, including Linear-CART, K-Nearest Neighbours–CART (KNN-CART), and a novel Feedforward Neural Network (FFNN) combined with Random Forest–CART (RF-CART). The proposed FFNN+RF-CART achieved 97.57% accuracy, 97.35% precision, 97.27% recall, and a 97.31% F1-score in cut classification. For price prediction, it delivered a Mean Absolute Error (MAE) of 0.1570, Mean Squared Error (MSE) of 0.1954, RMSE of 0.0004, and an  $R^2$  of 0.9870. This integrated framework offers a robust, data-driven approach for real-time diamond evaluation, enhancing transparency and efficiency in the gemmological industry.

**Keywords:** Price prediction, Diamond Classification, Feedforward Neural Network (FFNN), Random Forest-CART, Hybrid Models.

Received: 07-7-2025

Accepted: 11-8-2025

Published: 18-8-2025

**1. INTRODUCTION**

The global diamond industry remains one of the largest luxury markets worldwide as shown in Figure 1.1, valued at approximately USD 97.6 billion in 2024, with projections rising to USD 102 billion by 2025 and beyond [1]. Natural diamonds contribute the bulk of this value, while lab-grown diamonds, growing rapidly at around 15% per year, are reshaping demand dynamics [2]. The retail price per carat for high-quality diamonds has risen by roughly 6% in 2023, averaging USD 8,000, pointing to sustained appreciation

for premium stones. Despite steady long-term growth, the market has recently experienced a notable downturn. By early 2025, prices for natural diamonds had declined by approximately 26% from their 2022 peak, and lab-grown diamonds had plunged by up to 74% since 2020, largely due to technological advances and oversupply [3]. This sharp drop reflects weakening demand in key regions such as China and the U.S., along with declining marriage rates and increasing consumer shift to synthetic alternatives.

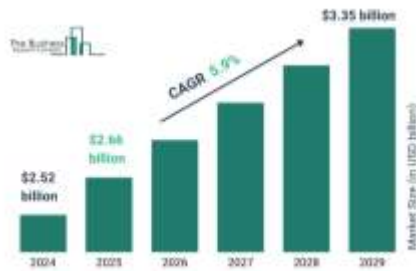


Fig. 1: Daimond demand global market.

Within the dataset commonly used for predictive modeling, which includes nearly 40,000 diamonds, descriptive statistics reveal an average price of around USD 3,927 per stone, with a standard deviation exceeding USD 3,980, and median price close to USD [4]. This underscores a highly skewed price distribution, driven by a mix of low-cost and very expensive diamonds. Moreover, features such as cut, color, and clarity exhibit strong correlations with price and quality outcomes, highlighting the importance of careful data exploration and preprocessing.

## 2. LITERATURE SURVEY

Ranglani et al. [5] proposed a unified machine learning framework that combined regression for price prediction and classification to distinguish natural diamonds from lab-grown specimens. They employed local linear regression to capture contextual relationships between features like carat, cut, clarity, and derived price, achieving a mean absolute error of \$554.32 on pricing and an  $F_1$ -score of 98.66% for origin classification. They provided interpretable insights into how feature impacts varied across subsets of data. The feature extraction remained limited to linear relationships and could not capture non-linear interdependencies. Amadavadi et al. [6] trained multiple regression algorithms including Random Forest, K-Nearest Neighbors, Decision Tree, and optimized model selection via performance metrics like  $R^2$ , MAE, and RMSE, using a Kaggle diamonds dataset of ~53,940 instances. Random Forest Regressor achieved the highest  $R^2$  (~0.9822) and lowest error values among compared models. They demonstrated clear comparative analysis across standard ML

models. They omitted advanced feature selection or deeper feature engineering and evaluated only raw categorical encodings, which limited performance gains. Basha et al. [7] benchmarked 23 machine learning models—both regression and classification—applying extensive hyperparameter tuning to optimize performance on price estimation tasks. They highlighted CatBoost, XGBoost, Random Forest, and LightGBM as top performers after tuning, and provided execution time comparisons alongside accuracy assessments. Their large benchmark offered valuable guidance on optimal model choice. The computational complexity soared with large model sets and hyperparameter grids, making deployment challenging in low-resource settings.

OuYang [8] investigated multiple linear regression, decision tree regression, and Random Forest regression models to predict diamond prices using evaluation metrics like MAE, RMSE, and adjusted  $R^2$ . Random Forest delivered the best predictive fit among the trio, with significantly lower RMSE and higher adjusted  $R^2$  values. They used simplicity and clear model comparison to inform use of data-driven pricing. The model ignored interactions among categorical quality factors (e.g. cut or clarity) and lacked any mechanism for encoding subtle qualitative distinctions. Reddy et al. [9] conducted a comprehensive study integrating classification and regression with hyperparameter optimization across dozens of models. They revealed consistent performance improvements post-tuning, with CatBoost and XGBoost retaining high accuracy even under cross-validated evaluation. The study offered actionable insights for pricing policy and inventory optimization in industry contexts. The feature extraction remained shallow and model interpretability suffered due to reliance on complex ensemble outputs without embedded explainable components. Fajar et al. [10] employed Random Forest Regressor, K-Nearest Neighbors (KNN) Regressor, and AdaBoost Regressor on a preprocessed Kaggle

diamond dataset of ~47,524 entries. Random Forest achieved the lowest test MSE (~132.73), RMSE (~11.52), and highest  $R^2$  (0.981), whereas KNN scored  $R^2$  of 0.964 and AdaBoost reached 0.884. They documented training duration (~13.73 s) and ranked models by predictive error. Their feature representation remained generic without explicit feature extraction or interaction modeling, limiting nuanced performance.

Bhadula et al. [11] proposed an explainable artificial intelligence (XAI) regression model for gold price forecasting based on economic indicators. The study employed model-agnostic explanation techniques to interpret the regression output, combining traditional regressors (e.g. random forest, linear models) with post hoc interpretation layers. They emphasized transparency by quantifying feature contributions and residual error margins, providing deeper insight into variable impacts. The model focused solely on post hoc interpretability without integrating explainable mechanisms within feature extraction, thus limiting understanding of latent feature interactions across time-series data. Kigo et al. [12] assessed multiple supervised algorithms including Random Forest, XGBoost, Multi-Layer Perceptron (MLP), Boosted Decision Trees, and KNN for both regression and classification tasks, after median imputation, outlier removal, equal-width binning for cut, and correlation-based feature selection. Random Forest regressor delivered RMSE of 523.50 and  $R^2$  of 0.985, while classifier achieved perfect AUC (1.00). MLP regressed with RMSE ~563.74 and  $R^2$  ~0.98, XGBoost slightly underperformed. Their classification of cut through coarse binning erased fine distinctions between adjacent quality levels. Acharya et al. [13] conducted a study on diamond price prediction using data mining techniques. They downloaded a Kaggle dataset of 53,940 diamonds with ten features and applied three regressors—Random Forest Regressor (RFR), K-Neighbors Regressor (KNN), and Decision Tree Regressor (DTR)—

alongside Particle Swarm Optimization (PSO) to fine-tune models. RFR achieved the highest  $R^2$  ( $\approx 0.9822$ ), followed closely by PSO, KNN, and DTR. The evaluation included MAE and RMSE to assess accuracy. The approach relied on standard regressors without explicit feature extraction or selection mechanisms to improve predictive robustness.

Ramírez et al. [14] introduced Extreme Learning Machine (ELM) variants—including standard, regularized, and weighted ELM—to classify diamond value into four price ranges. They handled classification into categories such as <US 500, 500–1000, 1000–1500, and >1500, and emphasized both classification accuracy (best was 0.8375) and model runtime (~109 s). Their method offered fast training due to the single-layer architecture of ELM. The classification results remained moderate and lacked fine-grained price prediction due to coarse value-range categorization, reducing precision in borderline cases. Mahanama et al. [15] studied machine learning techniques to assess prices of faceted blue sapphire gemstones. They collected real-world gemstone data and applied regressors such as Support Vector Regression (SVR), Random Forest Regressor (RFR), and Decision Tree Regressor (DTR) trained on features including color grade, carat, clarity indicators, and cut proportions. They highlighted how feature importance ranking aligned with domain knowledge of gemstone valuation. Their feature selection remained manual and basic, lacking automated extraction of latent interactions between numeric attributes. Swain et al. [16] presented GemInsight, a Random Forest (RF) based system for diamond quality forecasting, focusing on prediction of cut grade from tabular features. They preprocessed a dataset of diamond attributes, including polish, symmetry, carat, clarity, color, depth, and table percentages. They proposed domain-specific adjustments of tree depth and subtree pruning to enhance generalization. They operated directly on raw features without deep feature embedding, which limited capacity to model higher-order non-linear interactions.

Sun et al. [17] proposed a damage evaluation and property prediction framework for diamond tools under thermal stress using a small dataset of 2,870 scanning electron microscopy (SEM) images. Researchers produced samples with varying levels of heat-induced degradation, then applied a modified VGG16 convolutional neural network employing transfer learning and data augmentation strategies. They trained both classification models (identifying damage level categories) and regression models (predicting properties such as toughness index, flexural strength loss, relative density, and Rockwell hardness) and visualized damage zones with Grad-CAM saliency maps. The model depended heavily on image-level features without integrating semantic feature extraction tied to material microstructure, limiting interpretability and cross-sample generalization. Ghute and Korde [18] introduced an efficient machine learning (ML) algorithm for future gold price prediction using historical economic indicators and time-series features. They applied Random Forest Regressor (RFR) to forecast next-day gold closing price based on daily open, high, low, and volume data. They emphasized computational efficiency and scalability for financial stakeholders. The feature set remained limited to standard market metrics and omitted advanced feature extraction or latent time-series embeddings, restricting deeper interaction modeling.

### 3. PROPOSED SYSTEM

The proposed hybrid framework, illustrated in Fig.2z, introduces a unique combination of FFNN-based feature extraction integrated with multiple CART variants—namely Linear-CART, KNN-CART, and RF-CART—supported by SVM-SMOTE for data balancing.

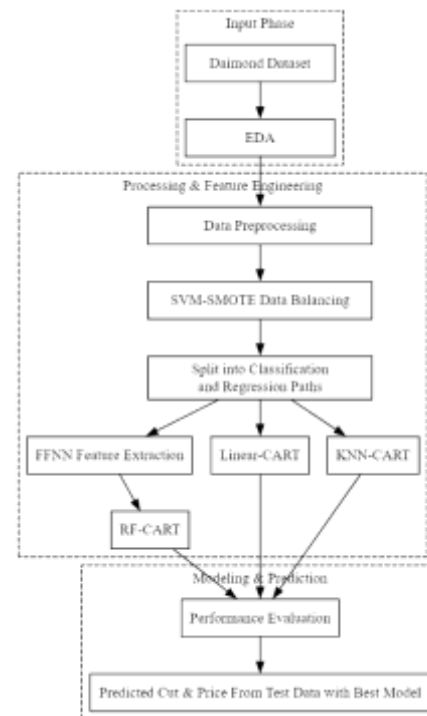


Fig. 2: Proposed system architecture.

#### 3.1 FFNN Feature Extraction with RF-CART

The FFNN Feature Extraction with RF-CART method combines the deep feature learning capability of a FFNN with the ensemble-based decision-making of Random Forest-based CART models for improved prediction in both classification and regression tasks. Initially, the FFNN processes the structured input features through multiple dense layers, where each layer progressively transforms the data into more abstract and task-relevant representations. The output from the penultimate layer is treated as the deep-learned feature set. These learned features are then fed into two separate Random Forest models — one classifier and one regressor — for final prediction. The classifier predicts discrete categories such as diamond cut quality, while the regressor estimates continuous values like price. This hybrid method leverages the powerful representation learning of FFNNs and the robustness and interpretability of RF-CART, offering improved generalization on complex structured data.

##### 3.1.1 FFNN Feature Extraction

The FFNN as shown in Fig. 3 approach for feature extraction is highly advantageous due

to its ability to model complex nonlinear relationships in both classification and regression tasks.

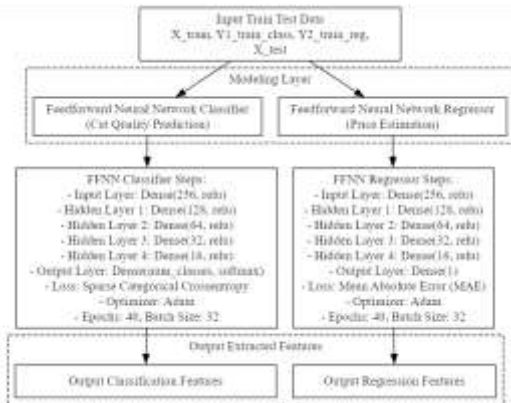


Fig. 3: FFNN feature extraction flowchart.

This method is particularly effective when the input dataset contains application-specific attributes (e.g., carat, color, clarity), where hierarchical representations are beneficial. The FFNN architecture extracts rich feature sets by progressively transforming inputs through multiple layers, each learning increasingly abstract patterns. This layered transformation allows for robust generalization and accurate predictions in both cut classification and price estimation. The same input dataset is efficiently reused for both tasks, reducing preprocessing redundancy and enabling simultaneous multi-objective learning.

### 3.1.2 RF-CART

The RF-CART as shown in Fig. 4 method is particularly well-suited for applications where both classification and regression tasks need to be performed using the same dataset. In this case, the dataset includes features that influence both the categorical output (cut quality of diamonds) and the numerical output (diamond price). The method benefits from its ability to automatically handle non-linearity, interactions among features, and overfitting through ensembling.

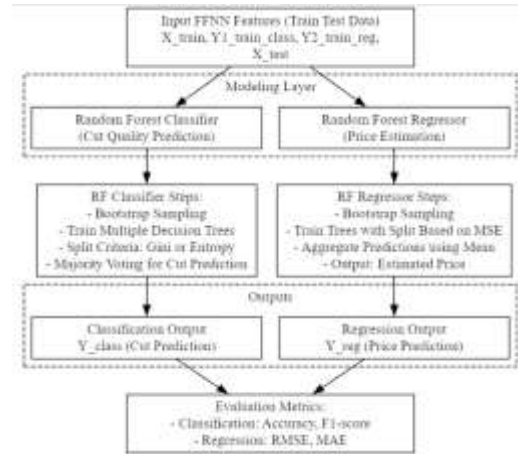


Fig. 4: RF-CART flowchart.

It is application-specific in that the nature and quality of input features greatly impact the model's predictive capability, and optimal results are achieved when the features are contextually relevant to the classification and regression targets.

## 4. RESULTS AND DISCUSSION

Fig. 5 shows histogram and KDE plot, where the distribution of diamond weights (carat). Most diamonds are under 1.5 carats, with peaks between 0.25 and 1.0 carats. Few diamonds exceed 2.5 carats, indicating that smaller diamonds are more common in the dataset. Fig. 6 shows count plot, which reveals that all five cut categories are nearly equally represented, each with around 21,000 samples. This balance is beneficial for training machine learning models, ensuring that no cut category is underrepresented or overfitted.

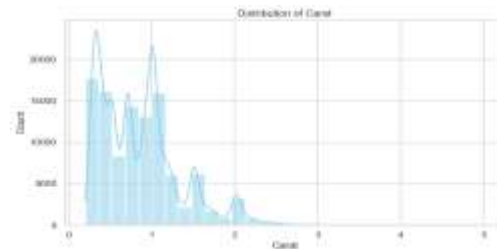


Fig. 5: Distribution of Carat in Dataset.

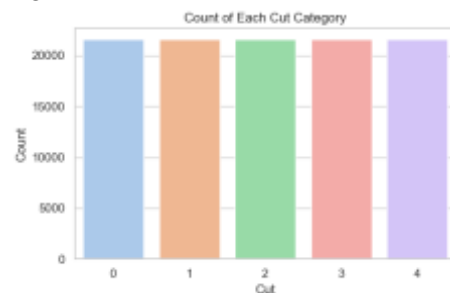


Fig. 6: Count plot of Cut.

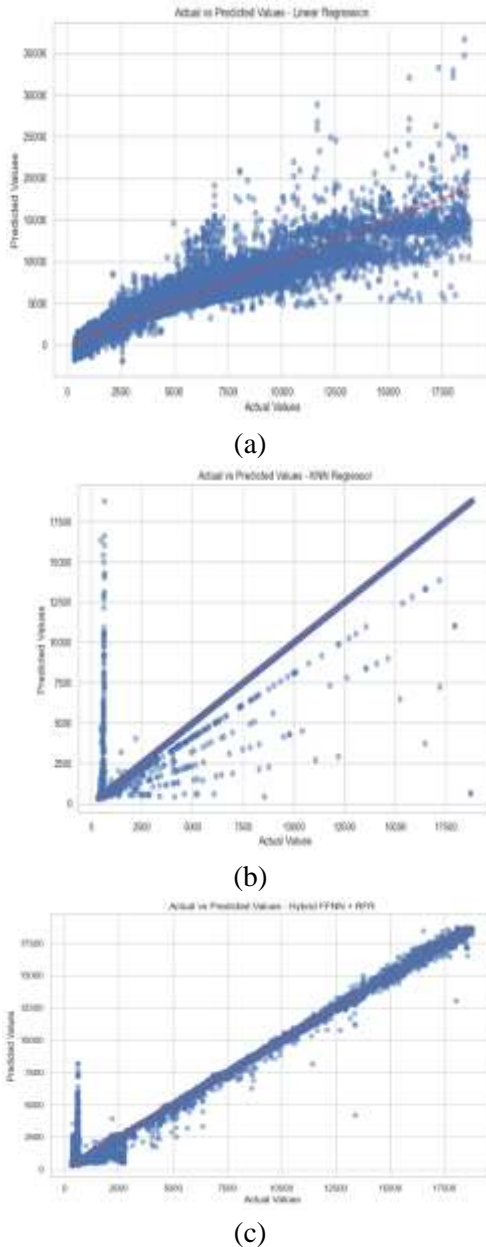


Fig. 7: Scatter Plot of Regression Models. (a) Linear-R. (b) KNN-R. (c) FFNN-RF-R.

Fig. 7 shows Scatter Plot of Regression Models. (a) Linear Regression: The scatter plot shows predicted vs. actual prices. Points are broadly aligned along the red 45° line but with noticeable spread and underestimation at higher values. This indicates Linear Regression captures general trends but struggles with complex relationships. (b) KNN Regressor: The predictions show a clustered trend, especially around lower values. While some predictions follow the actual values line, many deviate significantly, forming vertical

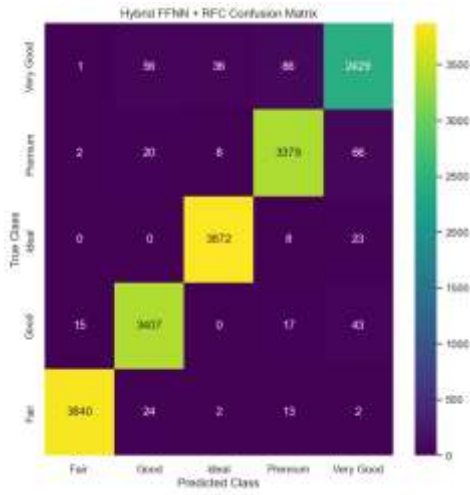
streaks. This behavior is typical of KNN when it overfits or fails to generalize(c) Hybrid FFNN + RFR: This hybrid model gives the best result. The predicted values align very closely with the actual values along the diagonal line with minimal scatter, indicating excellent prediction accuracy and generalization capability. It outperforms both Linear and KNN models significantly



(a)



(b)



(c)

Fig. 8: Confusion Matrices of Classification Models. (a) Linear-C. (b) KNN-C. (c) FFNN RF-C.

Fig. 8 presents the confusion matrices for three classification models: (a) Linear (Logistic) Classifier, (b) KNN Classifier, and (c) Hybrid FFNN + RFC. The Linear model, shown in Fig. 8(a), exhibits poor performance with an overall accuracy of 32.82%, correctly classifying Ideal most frequently (4083 true positives) but failing to identify Good, Premium, and Very Good classes effectively, with true positives of 0, 0, and 1 respectively, and substantial misclassifications such as 2456 Fair instances labeled as Ideal. The KNN model in Fig. 8(b) shows improved results with 58.53% accuracy, achieving more balanced class identification—Fair (3277), Good (2893), Ideal (2192), Premium (2288),

and Very Good (1964)—though notable errors remain, such as 306 Fair misclassified as Good and 661 Very Good as Premium. The Hybrid FFNN + RFC model, depicted in Fig. 8(c), delivers superior results with 97.57% accuracy and excellent class separation, recording high true positive counts for Fair (3840), Good (3407), Ideal (3872), Premium (3379), and Very Good (2429), with minimal misclassifications, such as 15 Fair as Good and 66 Premium as Very Good, highlighting its robustness and precision in predicting all diamond cut categories. Table. 1 summarizes the performance metrics of three models: Linear Regression, KNN Regressor, and a Hybrid FFNN + Random Forest Regressor (RFR). The Linear Regression model performs the weakest, with a high Mean Absolute Error (MAE) of 0.8979, Mean Squared Error (MSE) of 2.0251, and a low R<sup>2</sup> Score of 0.8655, indicating relatively poor predictive accuracy. In contrast, the KNN Regressor significantly improves performance, achieving a lower MAE of 0.0735, MSE of 0.4424, RMSE of 0.0007, and an excellent R<sup>2</sup> Score of 0.9706. The best results are obtained from the Hybrid FFNN + RFR model, with an MAE of 0.1570, the lowest MSE of 0.1954, RMSE of 0.0004, and the highest R<sup>2</sup> Score of 0.9870, highlighting its superior accuracy and reliability in predicting diamond price and cut with minimal error.

Table. 1: Performance comparison table for the diamond price prediction models.

Model	MAE	MSE	RMSE	R <sup>2</sup> Score
Linear Regression	0.8979	2.0251	0.0014	0.8655
KNN Regressor	0.0735	0.4424	0.0007	0.9706
Hybrid FFNN + RFR	0.1570	0.1954	0.0004	0.9870

Table 2: Performance comparison table for the diamond cut classification task.

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Linear Classifier	32.824	33.184	32.823	19.316
KNN Classifier	58.531	58.382	58.531	58.301
Hybrid FFNN + RFC	97.568	97.348	97.274	97.309

Table. 2 compares the performance of three algorithms for predicting diamond cut categories. The Linear (Logistic) Classifier performs the worst, with an accuracy of 32.824% and a low F1-score of 19.316%, showing poor balance and generalization across classes. The KNN Classifier performs better, achieving 58.531% accuracy with precision, recall, and F1-score all around 58%, indicating more consistent predictions. The Hybrid FFNN + Random Forest Classifier (RFC) delivers the best results by far, with 97.568% accuracy, 97.348% precision, 97.274% recall, and a 97.309% F1-score, clearly outperforming the other models and proving highly effective and reliable for predicting diamond quality categories.

## 5. CONCLUSION

The integration of FFNN and Random Forest models proves highly effective for both regression and classification tasks in diamond price and cut prediction tasks, outperforming individual traditional models across all evaluation metrics. The comparative evaluation of regression models highlights the superior performance of the Hybrid FFNN + RFR model. It achieved the lowest Mean Absolute Error (MAE) of 0.1570, Mean Squared Error (MSE) of 0.1954, and Root Mean Squared Error (RMSE) of 0.0004, indicating highly accurate predictions. Furthermore, its  $R^2$  Score of 0.9870 confirms excellent model fit and predictive capability. In contrast, the Linear Regression model demonstrated poor accuracy with an MAE of 0.8979, MSE of 2.0251, RMSE of 0.0014, and  $R^2$  Score of 0.8655, which shows a less reliable performance. The KNN Regressor offered a middle ground with MAE = 0.0735,

MSE = 0.4424, RMSE = 0.0007, and  $R^2$  Score = 0.9706, showing good but slightly less efficient performance than the hybrid model. In the classification scenario, the Hybrid FFNN + RFC model significantly outperformed traditional methods with Accuracy of 97.568%, Precision of 97.348%, Recall of 97.274%, and F1-Score of 97.309%, reflecting strong generalization and balanced learning. On the other hand, the Linear Classifier yielded Accuracy = 32.824% and F1-Score = 19.316%, highlighting weak classification ability. The KNN Classifier improved the performance with Accuracy = 58.531% and F1-Score = 58.301%, yet it was far from the robustness shown by the hybrid model.

## REFERENCES

- [1] April, Brian Kagiso, Lilian Oluoch, and O. Olawale Awe. "Novel Applications of Bayesian Additive Regression Model for Predicting Diamond Prices: A Comparative Study of Tree-Based Ensemble Techniques." In Practical Statistical Learning and Data Science Methods: Case Studies from LISA 2020 Global Network, USA, pp. 127-164. Cham: Springer Nature Switzerland, 2024.
- [2] Ranglani, Hardev. "Integrating Machine Learning for Diamond Price Prediction and Distinguishing Natural Diamonds from Lab Grown: A Unified Approach." In CS & IT Conference Proceedings, vol. 15, no. 1. CS & IT Conference Proceedings, 2025.
- [3] Zhou, Guanyun, Jianping Chen, Wentong An, Chang Liu, and Wei Li.

- "Three-dimensional mineral prospectivity mapping based on natural language processing and random forests: A case study of the Xiyu diamond deposit, China." *Ore Geology Reviews* 169 (2024): 106082.
- [4] Wahyusari, Retno, and Zahara Nabila. "Predicting Smartphone Prices Based on Key Features Using Random Forest and Gradient Boosting Algorithms in a Data Mining Framework." *International Journal for Applied Information Management* 5, no. 2 (2025): 73-85.
- [5] Ranglani, Hardev. "Integrating Machine Learning for Diamond Price Prediction and Distinguishing Natural Diamonds from Lab Grown: A Unified Approach." In *CS & IT Conference Proceedings*, vol. 15, no. 1. *CS & IT Conference Proceedings*, 2025.
- [6] Amadavadi, Kapil, Rashmi Rane, and Ruhi Patankar. "Diamond Price Prediction using Machine Learning Techniques." In *2024 5th International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 1994-2000. IEEE, 2024.
- [7] Basha, Md Shaik Amzad, Peerzadah Mohammad Oveis, R. Pavan Kumar Raju, and M. Martha Sucharitha. "Gems of Prediction: From Clarity to Carats-Unveiling Diamond Prices with Machine Learning in Waikato Environment for Knowledge Analysis." In *2024 International Conference on Distributed Computing and Optimization Techniques (ICDCOT)*, pp. 1-7. IEEE, 2024.
- [8] OuYang, Zhe. "Research on the diamond price prediction based on linear regression, decision tree and random forest." *Highlights in Business, Economics and Management* 24 (2024): 248-257.
- [9] Reddy, RN Bharath, KL Chandra Lekha, and Ongole Gandhi. "Enhancing Predictive Modeling of Diamond Prices using Machine Learning and Meta-Ensemble Techniques." In *2024 2nd International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC)*, pp. 275-280. IEEE, 2024.
- [10] Fajar, Fawzan Ibnu, Dian Sa'adillah Maylawati, Wildan Budiawan Zulfikar, Agung Wahana, Nur Lukman, and Muhammad Ali Ramdhani. "Random Forest, K-Nearest Neighbors, and Adaboost Regressor in Diamond Price Forecasting." In *2024 12th International Conference on Cyber and IT Service Management (CITSM)*, pp. 1-7. IEEE, 2024.
- [11] Bhadula, Shivam, Devashish Kartik, and Deepa Gupta. "An explainable AI regression model for gold price prediction." In *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, pp. 1-7. IEEE, 2024.
- [12] Kigo, Samuel Njoroge, Evans Otieno Omondi, and Bernard Oguna Omolo. "Assessing predictive performance of supervised machine learning algorithms for a diamond pricing model." *Scientific Reports* 13, no. 1 (2023): 17315.
- [13] Acharya, Kalpa Nigam, and Manas Ranjan Senapati. "Diamond Price Prediction Using Data Mining Techniques." In *International Conference on Research in Management & Technovation*, pp. 71-85. Singapore: Springer Nature Singapore, 2023.
- [14] Ramírez, José, David Zabala-Blanco, Roberto Ahumada-García, Juan Pablo Rivelli Malc6, Ali Dehghan Firoozabadi, and Marco Flores-Calero. "Extreme learning machines for predict the diamond price range."

- In 2023 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), pp. 1-6. IEEE, 2023.
- [15] Mahanama, Thidini, Vasana Chandrasekara, and Deershika Ranatunga. "Machine Learning Techniques to Assess Faceted Blue Sapphire Gemstone Prices." In 2024 International Research Conference on Smart Computing and Systems Engineering (SCSE), vol. 7, pp. 1-6. IEEE, 2024.
- [16] Swain, Debabrata, Avani Chaturvedi, and Vishesh Rathi. "GemInsight: Unleashing Random Forest for Diamond Quality Forecasting." In 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA), pp. 1-4. IEEE, 2023.
- [17] Sun, Wucheng, Hui Gao, Yuxiang Chen, Zhiming Wang, Longchen Duan, Songcheng Tan, and Xiaohong Fang. "Thermal failure of diamond tools indicated by diamond degradation: Damage evaluation and property prediction on small image datasets." *Engineering Applications of Artificial Intelligence* 119 (2023): 105800.
- [18] Ghute, Minal, and Mridula Korde. "Efficient machine learning algorithm for future gold price prediction." In 2023 International conference on inventive computation technologies (ICICT), pp. 216-220. IEEE, 2023.