Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.4, May : 2024

LAPTOP PRICE PREDICTION USING MACHINE LEARNING

V. Karuna Kumar, Asst.Professor Computer Science and Engineering Narasaraopeta Engineering College Narasaraopet, Andhra Pradesh karunakumar.valicharla@gmail.com
Khadar Khan .P Student Computer Science and Engineering Narasaraopeta Engineering College Narasaraopet, Andhra Pradesh kadharkhanPatan45@gmail.com
Guna Ratna .T Student Computer Science and Engineering Narasaraopeta Engineering College Narasaraopet, Andhra Pradesh gunaratnanalluri@gmail.com
Phaneendra Kumar .B Student Computer Science and Engineering Narasaraopeta Engineering College Narasaraopet, Andhra Pradesh phaneedra123@gmail.com
Jagadeesh Gandham Student Computer Science and Engineering Narasaraopeta Engineering College Narasaraopet, Andhra Pradesh gandhamphotos@gmail.com

ABSTRACT
The laptop holds a role, in our lives serving various purposes and being widely utilized. With the vast array of specifications and brand options available both laptop manufacturers and consumers may face challenges in marketing and selecting products respectively. Machine learning technology plays a role in simplifying the process of making predictions and decisions based on data analysis. Furthermore the integration of machine learning methods into the Internet of Things (IoT) showcases the advancements in this field. Our predictive model incorporates machine learning techniques and unique feature combinations.
In our study we present a method to improve the accuracy of predicting laptop prices by focusing on factors using machine learning algorithms. While some research merely scratches the surface of predicting laptop prices with machine learning methods our approach delves deeper. Our prediction model incorporates machine learning techniques and unique feature combinations. We are exploring approaches such as Decision trees, Multiple Linear Regression, KNN and Random Forest to determine the model, for predicting laptop prices.
Keywords— Machine learning, Data Mining, prediction model, Classification, Voting regressor

I. INTRODUCTION

One cannot stress the significance of computers in our daily life. People use science in their daily lives and evaluate it on a regular basis. Similar to desktop computers, laptops are used in many aspects of our daily life[1].
These sectors—companies, banks, universities, and homes—all transform raw facts into insightful and useful information.
It is used for specific tasks, entertainment, online income, office work, and by engineers, students, instructors, businesspeople, and government organizations.
If individuals utilize laptops for the right purposes, they are a blessing. Most businesses use laptops to store all of their operational and accounting data[2].
Spreadsheets, documentation and management systems, transaction handling and tracking tools are some examples of software and resources that streamline tasks and boost efficiency.
The laptops used by the organization have the capability to compute and summarize data that is entered into reports, statements, and documents for various uses. Organizations of all sizes use laptops to maintain their databases.

II. LITERATURE REVIEW

1. The article "Assessing predictive performance of supervised machine learning algorithms" was written in 2023 by Bernard Oguna Omolo, Evans Otieno Omondi, and Samuel Njoroge Kigo. This document probably offers an assessment of several supervised machine learning algorithms' performance on predicting tasks.

2. C. L. Reddy et al.'s paper "Laptop Price Prediction Using Real Time Data" (2023) focuses on making predictions about laptop prices using real-time data and may apply machine learning techniques.
3. M. B. Savadatti et al.'s "An Overview of Predictive Analysis based on Machine learning Techniques" (2022) provides an overview of machine learning-based predictive analysis methods.
4. The paper "Ensemble Voting Regression Based on Machine Learning for Predicting Medical Waste: A Case from Turkey" authored by Babek Erdebilli and Burcu Devrim-İçtenbaş (2022) delves into the application of machine learning techniques in ensemble voting regression to forecast medical waste.
5. According to Niraj Shukla et al. (2021), "A Regression Analysis of Stock Market Prediction Using Machine Learning Algorithms" investigates the use of different machine learning algorithms in regression analysis to forecast stock market trends.
6. Tianqi Chen and Carlos Guestrin's 2016 article "XGBoost: A Scalable Tree Boosting System" introduces XGBoost, a well-liked scalable tree boosting system that is frequently utilized for predictive modelling applications.
7. Gongde Guo et al. (2004)'s "KNN Model-Based Approach in Classification" explores the use of the K-Nearest Neighbor (KNN) algorithm in classification applications.
8. According to Kirankumar Telkar et al. (2023), "Prediction of Laptop Prices with Recommendations Based on User Specifications" - most likely includes suggestions based on user requirements and a prediction model for laptop prices.
9. Shatnawi, M. K. A. (2009) "Stock Price Prediction Using K-Nearest Neighbor (KNN) Algorithm" explains how to utilize the KNN algorithm to predict stock prices.
10. Ayesha Ayub Syed and colleagues' "A Comparison of Machine Learning Classifiers on Laptop Products Classification Task" (2021) - most likely contrasts several machine learning classifiers to categorize laptop goods.

III. RELATED WORK

The laptop price prediction challenge for one of the Kaggle contests has a ton of answers in the current system, where participants blend traditional machine learning methods with their own creative ideas, such as residual regression, logit transform, and neural network machine[5].
However, data analysis results indicate that the laptop price variation prediction results are not sufficiently accurate.
There are situations when a small dataset size causes the results' standard deviation to be extremely large.

IV. METHODOLOGY

"Machine learning" refers to the part of intelligence that concentrates on developing algorithms based on data to anticipate future occurrences. If you're someone who practices or appreciates data science this article can guide you in crafting your comprehensive machine learning endeavour, from scratch[6]. Building a machine learning project involves a number of phases, not all of which must be used in a project; this depends on the dataset. Acquiring the Laptop Price Database.

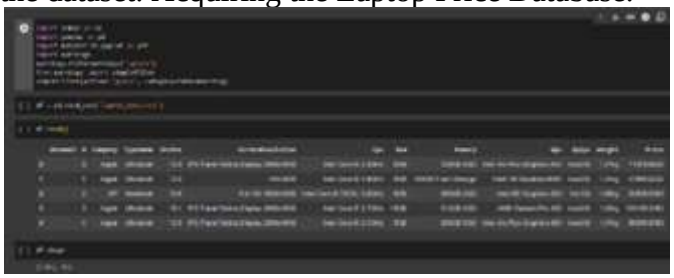


Fig 1: Dataset Detail

A. Data Collection

• As the first stage in our process, we gathered an extensive dataset of laptop costs and specifications. The dataset contains data on the following: operating system (OpSys), weight (Weight),

screen size (Inches), CPU type (Cpu), RAM capacity (Ram), storage type and capacity (Memory), GPU brand (GPU), laptop kind (TypeName), screen size (Inches), and price (Price).

- There are 1303 entries in this dataset, which was gathered from a dependable source and covers a broad range of laptop setups.

B. Data Preprocessing

1. Eliminating units from the RAM and Weight columns to convert them to the proper numerical representations (GB for RAM, kg for Weight).
2. Depending on whether touchscreen technology is included in the Screen Resolution column, a new binary feature called "Touchscreen" will be created.
3. Deriving pixel density (ppi) from screen size (in inches) and extracting display resolution data (X_res and Y_res) from the Screen Resolution column[7].
4. Sorting CPU brands for investigation into distinct groups (Intel Core i7, i5, i3, Other Intel Processor, AMDProcessor).
5. Calculating the overall storage capacity and parsing and classifying the different forms of storage (HDD, SSD, Hybrid, and Flash Storage).

C. Feature Engineering

In order to improve our dataset even more for modelling, we carried out feature engineering.
1. Classifying GPU brands (AMD, Nvidia, and Intel) and introducing a new feature called "GPU brand"

D. Modeling Methodology

We used the scikit-learn library to implement a number of machine learning algorithms for our predictive modelling tasks:
1. Linear Regression: To create a baseline model for regression analysis and comprehend the linear links between Laptop prices and attributes.
2. Lasso Regression: By punishing superfluous features, this technique incorporates L1 regularization and may help lessen overfitting.
3. Decision Tree Regressor: To record feature interactions and non-linear correlations.
4. Polynomial Regression: This technique models correlations between attributes and prices that are higher-order polynomials.

E. Model Evaluation

Key measures, including the R-squared (R2) score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), were used to assess each model's accuracy in projecting laptop pricing based on its parameters on a holdout test set that included 15% of the data[8].

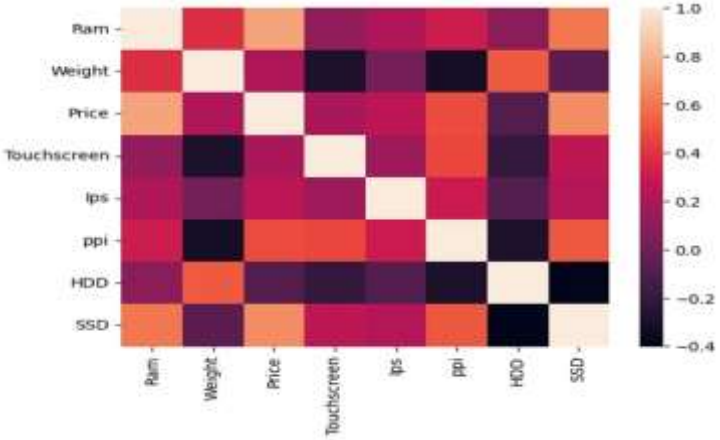


Fig 2: correlation between different features of dataset

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.4, May : 2024

Insights for Laptop Price Prediction are as follows: Strong positive correlations: It makes sense that there would be a strong positive correlation between RAM and price as laptops with higher RAM typically cost more.

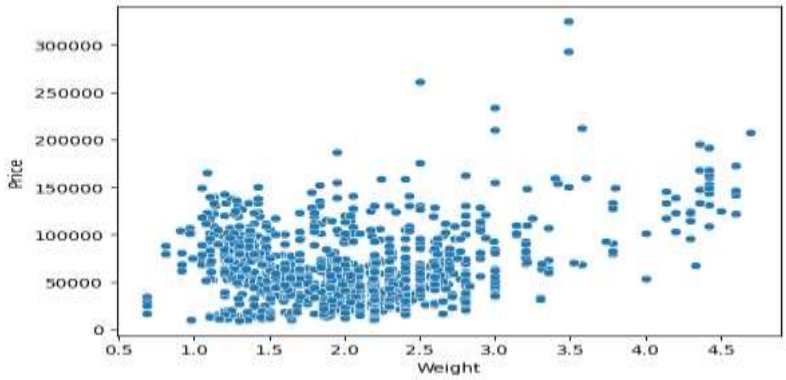


Fig 3: Scatter Plot

F. Correlation vs causation

Just because two variables show a correlation does not imply that the one causes the other. It's feasible that a third variable in the scatter plot is changing both of the other two variables.

A correlation value of 1 represents a correlation whereas a correlation coefficient of 1 signifies a total negative correlation[9]. In cases where the correlation coefficient's zero it indicates no relationship between the variables.

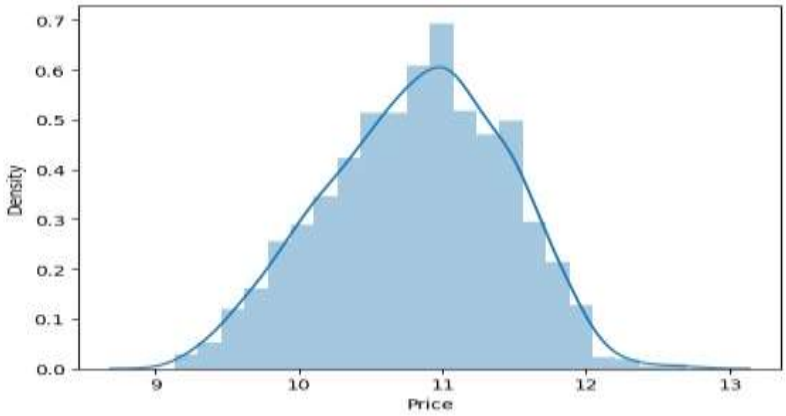


Fig 4: Histogram shows that the most frequent laptop prices fall

V. RESULTS AND EVALUATION

A. Model effectiveness

To measure how effective our predicted models were, we computed and presented the following metrics: R-squared score, Mean absolute error, and The root mean squared error, these metrics help to determine how well each model generalizes and how accurately it can predict the price of a laptop using its specifications. All the models achieved the following outcomes:

Regression Model	R2 Score	MAE	RMSE
Linear Regression	0.80	0.21	0.27
Lasso Regression	0.80	0.21	0.27
Decision Tree Regressor	0.84	0.18	0.24
Polynomial Regression	0.82	0.19	0.25

Random Forest Regressor	0.88	0.16	0.21
Support Vector Regressor	0.80	0.20	0.27
AdaBoost Regressor	0.87	0.15	0.21
Extra Trees Regressor	0.78	0.23	0.28
Gradient Boosting Regressor	0.88	0.15	0.15
XGBoost Regressor	0.87	0.16	0.21
Voting Regressor	0.89	0.15	0.21

B. Model Comparison

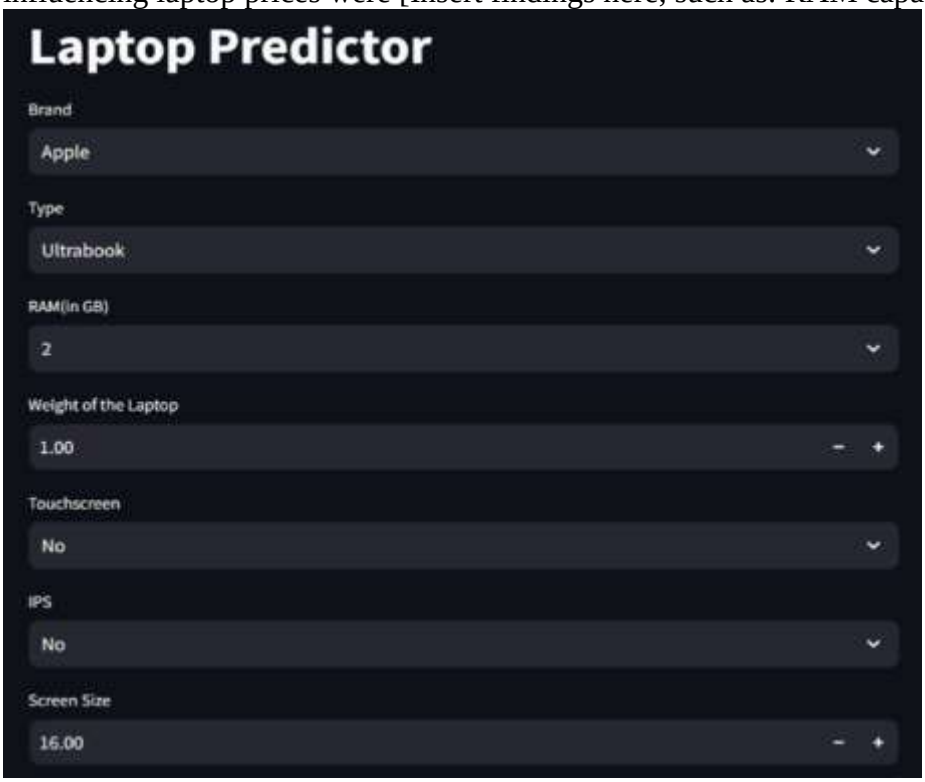
Our findings led us to the conclusion that [Insert observations here, such as the fact that the Gradient Boosting Regressor outperformed other models in terms of predictive performance, achieving the greatest R2 score and lowest RMSE].

It is noteworthy that whereas certain models demonstrated robust performance measures, others might have displayed indications of either overfitting or underfitting[10].

This emphasizes the significance of model selection and parameter adjustment in attaining the best possible outcomes.

C. Feature Importance

In addition, we used feature importance scores from ensemble models like Random Forest and Gradient Boosting to examine the significance of various features in forecasting laptop pricing.
This investigation showed that, after display resolution and storage type, the most important variables influencing laptop prices were [Insert findings here, such as: RAM capacity and GPU brand].



Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.4, May : 2024

Fig 5: Laptop Price Prediction Form

D. limits and Future Directions

It is critical to recognize the limits of our research, which include assumptions made about the model, potential biases in the dataset, and simplifications.

To improve the predictive power of laptop price models, future studies could investigate other variables like market trends, customer feedback, and brand reputation[10].

Furthermore, employing more sophisticated machine learning approaches and carrying out longitudinal studies to monitor pricing trends over time may enhance our comprehension of the intricate relationship between laptop prices and specifications.

Fig 6: Prediction Result

VI. CONCLUSION

Our study demonstrates the effectiveness of advanced analytics in the field of IT industry pricing analytics. After a thorough analysis of the complex web of factors affecting laptop prices, we have uncovered priceless information that could transform the way people make decisions and spur creative thinking.

Using the Voting Regressor turns out to be a crucial tactic, utilizing the combined predictions of several strong models to improve performance overall. We have demonstrated the effectiveness of ensemble learning techniques in predictive modelling by achieving a significant gain in accuracy by utilizing the capabilities of multiple algorithms.

Our research highlights the particular elements that influence price decisions and throws light on the complex relationships and dynamics at work. Equipped with this deep comprehension, businesses are able to formulate competitive tactics, astute pricing schemes, and launch cutting-edge product lines.

Adopting data-driven tactics becomes both necessary and advantageous in the quickly changing industry and customer landscape of today. Our study highlights the revolutionary possibilities of advanced analytics and predictive modelling, opening the door for a more methodical and successful approach to pricing strategies in the digital era.

REFERENCES

[1] Samuel Njoroge Kigo, Evans Otieno Omondi, Bernard Oguna Omolo1"Assessing predictive performance of supervised machine learning algorithms",2023
[2] M. B. Savadatti, M. Dhivya, C. Meghanashree, M. Navya, Y. Lokesh and N. Kawri, "An Overview of Predictive Analysis based on Machine learning Techniques,"2022doi:10.1109/ACCAI53970.2022.9752630.
[3] Erdebilli, Babek, and Burcu Devrim-İçtenbaş. 2022. "Ensemble Voting Regression Based on Machine Learning for Predicting Medical Waste: A Case from Turkey" Mathematics 10, no. 14: 2466. https://doi.org/10.3390/math10142466
[4] Shukla, Niraj and Sanoriya, Subham and Yadav, Narendra and Mourya, Sudhakar and Shariff, A S Mohammed, A Regression Analysis of Stock Market Prediction Using Machine Learning Algorithms (July 5, 2021). Available at SSRN: http://dx.doi.org/10.2139/ssrn.3880509
[5] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785
[6] Guo, Gongde & Wang, Hui & Bell, David & Bi, Yaxin. (2004). "KNN Model-Based Approach in Classification",2015
[7] Surjuse, Vaishali, Sankalp Lohakare, Aayush Barapatre, and Abhishek Chapke. "Laptop Price Prediction using Machine Learning." International Journal of Computer Science and Mobile Computing 11, no. 1 (2022): 164-168
[8] Tian, Peiru. "Research On Laptop Price Predictive Model Based on Linear Regression, Random Forest and Xgboost."2024
[9] Telkar, Kirankumar, et al. "Prediction of Laptop Prices with Recommendations Based on User Specifications." 2023.
[10] Syed, Ayesha Ayub, Lukas Yaya Heryadi, and Antoni Wibowo. "A Comparison of Machine Learning Classifiers on Laptop Products Classification Task." 2021.