Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 7, April : 2024

LAPTOP PRICE PREDICTION USING MACHINE LEARNING

Anant Garg, Dept. of CSE (AIML), ABES Engineering College, AKTU University
Deepanshu Sharma, Dept. of CSE (AIML), ABES Engineering College, AKTU University
Ishan, Dept. of CSE (AIML), ABES Engineering College, AKTU University
Mrs. Parul Phoghat, Assistant Professor, Dept. of CSE (AIML), ABES Engineering College, AKTU University

Abstract
Accurate prediction of laptop prices is vital is the ever -changing landscape of technology, benefiting consumers, retailers, and manufacturers alike. This research focuses on leveraging Machine Learning algorithms to predict laptop prices, and develop robust models for laptop price prediction. Utilizing a diverse dataset containing various laptop specifications and historical pricing data, the study employs advanced machine learning algorithms, including Linear Regression, K-Nearest Neighbors, Decision Tree Regressor, Random Forest Regressor.
This research titled "Laptop Price Prediction" endeavors to develop a precise and dependable machine learning model capable of predicting laptop prices based on diverse features and specifications. The dynamic nature of technology and the vast variety of laptops available in the market make determining fair market values challenging. This research addresses this complexity by utilizing a diverse dataset that encompasses a wide range of laptop models, brands, and technical specifications. The proposed solution employs advanced machine learning algorithms, including regression and ensemble methods, to analyze the dataset and identify patterns that significantly influence laptop pricing. Feature engineering techniques are applied to extract pertinent information from the dataset, taking into account factors such as processor speed, RAM capacity, storage type, screen size, brand reputation, and other essential features. Moreover, this research aims at how a user-friendly interface or application can be created that allows users to input laptop specifications and receive an estimated price prediction.
Additionally, the study explores the impact of feature selection and engineering on model predictive capabilities. The outcomes not only aid consumers in making informed purchase decisions but also provide actionable insights for retailers and manufacturers to optimize pricing strategies and enhance competitiveness in the dynamic laptop market. In this research Random Forest Regressor (Supervised Machine Learning Algorithms) results with the least MAE (Mean Absolute Error) of 0.157 and highest R^2 (R-Squared) value of 0.887.

Keywords:
Laptop, Price Prediction, Laptop Features, One Hot Encoding, Supervised Machine Learning Algorithms, Linear Regression, K Nearest Neighbors, Decision Tree Regressor, Random Forest Regressor, Streamlit, Web Application.

I. Introduction

In today’s fast-paced technological landscape, laptops have become indispensable tools for both personal and professional use. Laptop is a small, portable personal computer (PC) with a screen & alphanumeric keyboard. It has many features or configurations like RAM (Random Access Memory), ROM (Read Only Memory), CPU(Central Processing Unit), Weight, Inches, GPU (Graphics Processing Unit), etc.
In India, the demand for laptops soared after the nationwide lockdown. This device is now used by many technical individuals. Before purchasing any laptop, people have specific requirements, such as the required RAM capacity, among other factors. Therefore, there is a need for a system that can provide an estimate of the expected cost for a laptop that meets the user's requirements.

With the abundance of options available in the market, consumers often grapple with the challenge of selecting a laptop that aligns with their needs while staying within budgetary constraints. Accurately predicting laptop prices is not only valuable for consumers seeking informed purchasing decisions but also for retailers and manufactures aiming to optimize pricing strategies and maintain competitiveness. Conventional pricing estimation methods may lack the precision and adaptability required to navigate the ever-evolving laptop market. Therefore, leveraging machine learning techniques presents a promising avenue to develop robust models for laptop price prediction.

Additionally, the study aims to delve into the key factors influencing laptop pricing, illuminating the underlying trends and patterns within the market. Through an exhaustive examination of various machine learning algorithms, including regression-based models, decision trees, and ensemble methods.

This research aims to transcend theoretical model development by creating a user-friendly interface. This interface serves as a practical tool for consumers, retailers, and manufacturers, allowing them to input laptop specifications and obtain reliable price predictions.

Flow of Logic:

1) Collect Data & Select Required Data
2) Prepare, Preprocess, Transform Data
3) Model is trained using training data
4) Model that forecasts prices with the highest accuracy rate will be chosen to power an application

II. Literature

The evolving landscape of consumer electronics has spurred research interest in predictive modeling for pricing strategies, with a particular focus on laptops. This review aims to provide insights into existing studies, emphasizing the current state of knowledge and identifying gaps that the project "Laptop Price Prediction" seeks to address. Our research aims to contribute to a more nuanced understanding of the factors influencing laptop prices and to provide actionable insights for industry stakeholders, ultimately advancing the field of price prediction.

2.1 Factors Influencing Laptop Prices: Numerous studies have explored the determinants of laptop prices, consistently highlighting processor specifications, RAM capacity, storage types, and brand reputation. These findings align with the proposed project's goal of unraveling the intricate interplay of these features.

2.2 Machine Learning in Pricing Strategies: Recent literature showcases the increasing integration of machine learning techniques in pricing strategies. Algorithms such as regression and ensemble methods have proven effective in predicting prices based on diverse datasets, offering valuable insights into market dynamics.

2.3 Relevant Studies on Predicting Electronic Device Prices: Several studies have specifically delved into predicting prices for electronic devices. Noteworthy research has focused on smartphones, where features like camera quality, storage capacity, and brand recognition significantly influence pricing. These studies have provided valuable insights into the nuanced relationship between device specifications and market value, offering a foundation for applying similar methodologies to the laptop market.

2.4 Prof. Vaishali Surjuse, Sankalp Lohakare, Aayush Barapatre, Abhishek Chapke [4] authored a paper introducing a Laptop Price Prediction System through the application of Supervised Machine Learning technique. The study employed Multiple Linear Regression as the chosen machine learning prediction method, achieving a predictive accuracy of 81%.

2.5 Prof. Prajna, Lekha Sri, Sahiti, Bhagyasri Teja, Arzoo [2] authored a paper introducing Laptop Price Prediction Using Machine Learning in which they discussed how we can use Multiple Linear Regression a supervised machine learning algorithm which offered 81% prediction precision can be used to predict the laptop prices.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 7, April : 2024

2.6 Listian discussed the effectiveness of a regression model based on the Decision Tree Regressor in predicting the prices of laptops. The findings suggested that this model outperformed multivariate regression and simple multiple regression in terms of precision. One key advantage of the Decision Tree Algorithm highlighted in the research is its suitability for handling datasets with a higher number of dimensions. Additionally, it was noted that Decision Tree models tend to be less susceptible to overfitting and underfitting compared to some other regression methods.

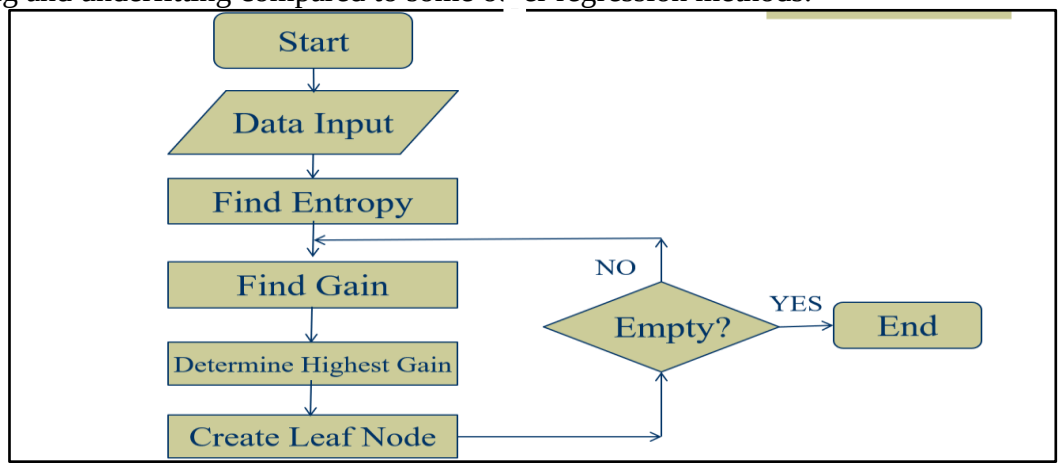

Figure 1: Flowchart of Decision Tree Algorithm

III. Proposed Methodology

The methodology for this research on "Laptop Price Prediction" is carefully structured to navigate through key phases of model development, data handling, and interface creation. Here's a detailed breakdown:

3.1 Data Collection: The initial step involves gathering a diverse and comprehensive dataset that encapsulates various laptop models, brands, and technical specifications. This dataset should be inclusive of crucial features such as processor speed, RAM capacity, storage type, screen size, and brand reputation. For this research we used the dataset available on Kaggle having several features which affects the price of a laptop.

```
In [10]: df.head()
```

Out[10]:

	Company	TypeName	Inches	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight	Price
0	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8GB	128GB SSD	Intel Iris Plus Graphics 640	macOS	1.37kg	71378.6832
1	Apple	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8GB	128GB Flash Storage	Intel HD Graphics 6000	macOS	1.34kg	47895.5232
2	HP	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8GB	256GB SSD	Intel HD Graphics 620	No OS	1.86kg	30636.0000
3	Apple	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16GB	512GB SSD	AMD Radeon Pro 455	macOS	1.83kg	135195.3360
4	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8GB	256GB SSD	Intel Iris Plus Graphics 650	macOS	1.37kg	96095.8080

Figure 2: Rows of Laptop Dataset with labeled features

3.2 Data Preprocessing: It is the concept of changing the raw data into a clean dataset. This includes addressing missing values, handling outliers, and ensuring consistency. Numerical features are normalized or standardized to maintain uniformity, and categorical variables are transformed into numerical representations through suitable encoding techniques.

3.3 Feature Engineering: [12] Feature engineering focuses on identifying and extracting relevant features that wield significant influence over laptop prices. It is a machine learning technique that transforms raw data into a more effective set of inputs, called features, for use in supervised and unsupervised learning. This may involve creating new features or combining existing ones to capture nuanced relationships. Techniques for feature selection are applied to retain the most informative variables.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 7, April : 2024

3.4 Exploratory Data Analysis: As we discussed there are many features of laptops, so we perform visual analysis on various features like Company, Weight, Operating System etc. EDA is a statistical approach for summarizing the main characteristics of data sets [12].The analysis is mainly done between the feature and target. Bar Plot & Scatter Plot are the main charts used for the visualization. Here we added Company vs Price Bar Plot & Weight vs Price Scatter Plot.

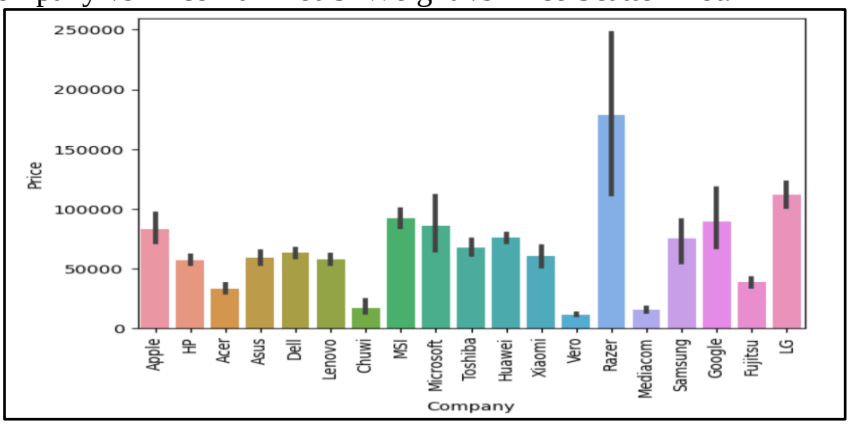

Figure 3: Company vs Price Bar Plot

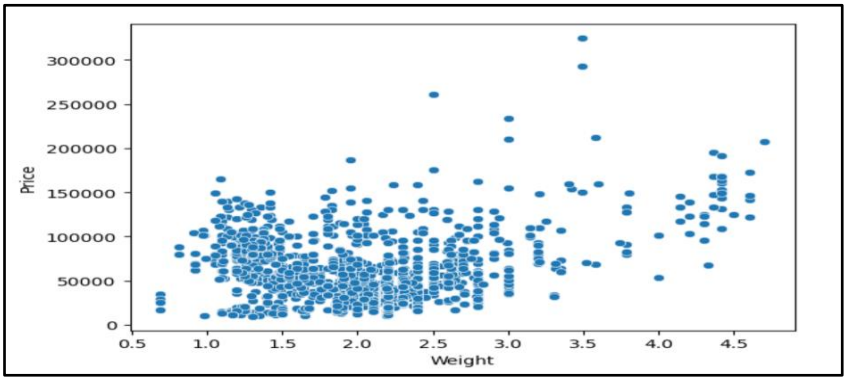

Figure 4: Weight vs Price Scatter Plot

3.5 Algorithm Selection: Various machine learning algorithms are evaluated to identify the most fitting for regression tasks. This includes popular choices such as linear regression, k nearest neighbor, decision trees, random forests, and ensemble methods. The selection process hinges on finding the algorithm that strikes the best balance between accuracy and generalization, as determined by performance metrics.

3.6 Model Training: The dataset is divided into training and validation sets for the model training phase. Fine-tuning of hyper parameters takes place to optimize the model's performance. Cross-validation techniques are implemented to ensure the model's robustness and prevent over-fitting.

3.7 Model Evaluation & Testing: The model's performance is rigorously evaluated using metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared. A crucial step involves comparing predicted prices against actual prices from the validation set to gauge accuracy and reliability.

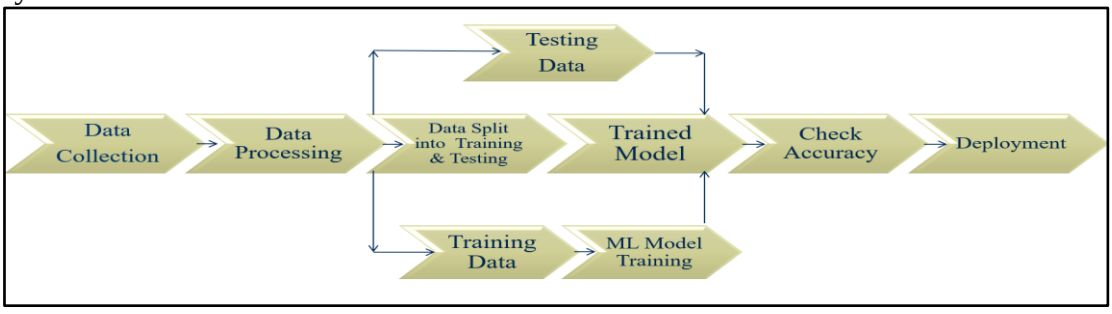

Figure 5: Flowchart of Machine Learning

3.8 User Interface Development: A user-friendly interface or application is designed and developed, allowing users to input laptop specifications and receive accurate price predictions. Accessibility is prioritized for diverse stakeholders, including consumers, retailers, and manufacturers. This can be directly done by using Python Libraries streamlit & pickle, or this can also be done with the help of HTML, CSS & JavaScript.

IV. Results

Mean Absolute Error (MAE): Mean Absolute Error (MAE) is a metric used to evaluate the accuracy of a predictive model by measuring the average absolute differences between predicted and actual values. It provides a simple and straightforward measure of how well the model is performing in terms of the absolute magnitude of errors.

R-squared (R^2): R-squared (R^2) quantifies the proportion of variance in the dependent variable (price) that is predictable from the independent variables, offering insights into the overall goodness of fit of the model.

This table representing the Mean Absolute Error (MAE) and R-Squared (R^2) values of different Machine Learning Algorithms that we have used and Random Forest Regressor is performing in best manner with least MAE of 0.15751604016 and high R^2 value of 0.88710207744.

ALGORITHMS	MAE	R^2
Linear Regression	0.21022212996	0.80732574309
K-Nearest Neighbors	0.19434951058	0.80145442655
Decision Tree Regressor	0.17825091972	0.84853479042
Random Forest Regressor	0.15751604016	0.88710207744

Table 1: Comparison of MAE & R^2 of Machine Learning Algorithms

Interpretation of Results: The findings of our research on predicting laptop prices using Machine Learning highlight the models' effectiveness in providing accurate price estimations. The Mean Absolute Error (MAE) values indicate that the models achieved predictions with relatively small errors, showcasing their practical utility. The R-squared values further support the notion that a significant proportion of the variance in laptop prices can be explained by the selected features. The scatter plots and residual plots visually demonstrate the alignment between predicted and actual prices, reinforcing the models' reliability. Overall, our models performed well in capturing the complex relationships between various laptop specifications and their market prices. Hence Random Forest Regressor is resulting to give less Mean Absolute Error (MAE) and more R-squared (R^2), Since this results in more suitable Machine Learning Algorithms.

Limitations: Despite the promising results, our study is not without limitations. One notable constraint is the potential for biases in the dataset, which may stem from variations in data collection methodologies across different sources. Moreover, the dynamic nature of the technology market introduces challenges in keeping the dataset current, potentially affecting the models' adaptability to rapidly changing trends. Additionally, assumptions about the linear relationships between features and prices may not hold universally, and the models may struggle with capturing subtle interactions or non-linear patterns.

Future Work: Future research endeavors should address these limitations and explore several areas for improvement and expansion. Firstly, incorporating real-time data updates could enhance the models' responsiveness to market changes. Exploring advanced modeling techniques, such as neural networks, may provide a more nuanced understanding of the intricate relationships within the laptop market. Additionally, considering sentiment analysis from user reviews and incorporating external

economic factors could further refine the models' predictive capabilities. Finally, a comparative analysis across different geographical markets and cultural contexts could unveil region-specific pricing dynamics, adding a layer of granularity to the models.

V. Conclusion

Key Findings and Implications: Our research on predicting laptop prices using Machine Learning yielded several key findings with significant implications. The machine learning models developed demonstrated commendable accuracy, as indicated by low Mean Absolute Error (MAE). The models effectively captured the relationships between diverse laptop specifications and their market prices, as evidenced by high R-squared values. Visualizations, including scatter plots and residual plots, reinforced the reliability of the models in predicting laptop prices.

Implications:

- **Informed Consumer Decisions**: Accurate price predictions empower consumers to make well-informed decisions when purchasing laptops. This is particularly crucial given the diverse range of laptops available, allowing consumers to match their needs with affordable options.
- **Optimized Manufacturing Strategies:** Manufacturers can leverage price predictions to optimize product offerings and pricing strategies. Understanding which features significantly influence pricing enables manufacturers to align their product development with consumer preferences and market trends.
- **Efficient Retail Operations**: Retailers stand to benefit from accurate price predictions by optimizing inventory management and pricing strategies. This ensures that laptops are priced competitively, leading to improved sales and customer satisfaction.
- **Market Insights**: The research provides valuable insights into market dynamics, showcasing the significance of various specifications in influencing laptop prices. This knowledge can guide stakeholders in adapting to changing consumer preferences and technological trends.

Reiterating the Importance: Predicting laptop prices is integral to fostering a more transparent, efficient, and consumer-eccentric laptop market. It plays a pivotal role in democratizing access to technology, enabling consumers across diverse demographics to make informed choices within their budget constraints. For manufacturers, the ability to predict prices facilitates strategic decision-making, ensuring that product offerings are not only technologically advanced but also competitively priced. Retailers benefit from optimized operations, creating a win-win situation where consumers find suitable options, manufacturers tailor products effectively, and retailers operate efficiently in a dynamic market. In essence, predicting laptop prices is a cornerstone for creating a thriving and inclusive ecosystem within the laptop industry, where stakeholders at every level can navigate the complexities of the market with confidence and strategic foresight.

References

[1]. Parmeshwar M, Komal N. (2023) Laptop Prediction & Comparison using Machine Learning. International Journal of Innovative Research in Technology.
[2]. Prof. B.Prajna, A.Lekha Sri, A.Sahiti, A.Bhagyasri Teja, A.Arzoo (2023) Laptop Price Prediction Using Machine Learning. International Journal of Novel Research and Development (IJNRD).
[3]. Cagliero, Luca, Jacopo Fior, and Paolo Garza. 2023. Shortlisting machine learning-based stock trading recommendations using candlestick pattern recognition.
[4]. Vaishali S, Sankalp L. (2022) Laptop Price Prediction using Machine Learning. International Journal of Computer Science and Mobile Computing.
[5]. Proceeding of the International Multiconference of Engineers and Computer Scientists 2021. A Comparison of Machine Learning Classifiers on Laptop Products Classification Task.

[6]. Ali, Syed Shahan, Muhammad Mubeen, and Adnan Hussain. 2018. Prediction of stock performance by using logistic regression model: Evidence from Pakistan Stock Exchange (PSX). Asian Journal of Empirical Research
[7]. Noor, K., & Jan, S. (2017). Vehicle Price Prediction System using Machine Learning Techniques. International Journal of Computer Applications, 167(9), 27-31.
[8]. Pudaruth, S. (2014). Predicting the price of used cars using machine learning techniques. Int. J. Inf. Comput. Technol, 4(7), 753-764.
[9]. Pandey M, Sharma VK. A decision tree algorithm pertaining to the student performance analysis and prediction. International Journal of Computer Applications. 2013 Jan 1;61(13)
[10]. Aminah Md Yusof and Ismail Syuhaida, "Multiple Regressions in Analyzing House Price Variations", *IBIMA Publishing Communications of the IBIMA*, vol. 2012, pp. 9, 2012.
[11]. Listiani , M. (2009). Support vector regression analysis for price prediction in a car leasing application (Doctoral dissertation, Master thesis, TU HamburgHarburg).
[12].www.researchgate.net/publication/50946368_Exploratory_data_analysis_in_the_context_of_data_mining_and_resampling