International Journal of Research in Engineering, Science and Management
Volume 7, Issue 3, March 2024
https://www.ijresm.com | ISSN (Online): 2581-5792

Laptop Price Estimation Using Machine Learning

Puvvada Pragnatha[1], Korada Sai Yaswanth Kumar[2], Vemala S. S. S. M. Vikas[3], Sai Pavan Gorle[4], Nookala Subrahmanya Sai[5*], Sangam Hemanth Sai[6]

[1]*Assistant Professor, Department of Computer Science and Information Technology, Lendi Institute of Engineering and Technology, Vizianagaram, India*
[2,3,4,5,6]*B.Tech. Student, Department of Computer Science and Information Technology, Lendi Institute of Engineering and Technology, Vizianagaram, India*

***Abstract*: Predicting the cost of a laptop is a crucial and significant undertaking, particularly when the device is being sent straight from the manufacturer to the electronic market or stores. Expert knowledge is necessary for accurate laptop price prediction because the price is typically influenced by a wide range of unique features and circumstances. In this project, we develop a Laptop Price Prediction Model aimed at aiding students and common individuals in making informed and budget-conscious decisions when purchasing laptops. Leveraging machine learning techniques, particularly the Random Forest algorithm, the model predicts laptop prices based on diverse features, ensuring accuracy in cost estimation. The web application, built using Streamlit, provides an intuitive interface for users to input laptop specifications, allowing them to receive real-time price predictions. This project addresses the practical need for budget-friendly laptop choices, enhancing accessibility and facilitating optimal decision-making in the realm of technology purchases. The streamlined web application ensures a user-friendly experience, making it a valuable tool for both students and the general public seeking reliable guidance in the complex landscape of laptop pricing.**

***Keywords*: Streamlit, Random Forest algorithm.**

1. Introduction

The Laptop Price Prediction System stands as an innovative solution poised to transform the laptop market's pricing dynamics. Utilizing cutting-edge machine learning algorithms, the system seeks to deliver precise and effective price predictions, empowering users to make well-informed decisions during laptop purchases. This comprehensive project spans multiple phases, including exploratory data analysis, feature engineering, machine learning modeling, web application development, and meticulous testing. The integration of Streamlit for the web app and the Random Forest algorithm for predictive modeling enhances the system's robustness. Our objective is to provide a user-friendly and accurate tool that equips consumers to navigate the dynamic landscape of laptop pricing with confidence.

2. Methodology

A. Data Cleaning

The attributes present in the dataset are Company, RAM, Type Name, Inches, Screen resolution, memory, OpSys, GPU, CPU, weight, price. There are no NULL values in the dataset. Weight and Ram columns are converted to numeric by removing the unit written after the value in order to make the data more suitable for numerical computations or visualizations.

B. Exploratory Data Analysis

An essential part of our project is the exploratory data analysis (EDA) phase, which entails a detailed examination of the dataset to comprehend its subtleties and connections. The purpose of this procedure, which includes Univariate, Bivariate, and Multivariate analyses, is to test hypotheses and validate assumptions. We execute preprocessing and feature engineering procedures to improve data quality, going over each feature one by one starting with the first column. The ultimate aim is to thoroughly prepare and clean the data, providing the groundwork for the creation of machine learning models that are both highly effective and broadly applicable.

C. Feature Engineering

Feature engineering and preprocessing are important steps in improving our laptop price prediction model. By addressing noisy columns such as Screen Resolution, CPU, RAM, and Memory, we are able to extract useful information that helps us to improve the dataset.

Effective feature engineering is integral to the refinement of our Laptop Price Prediction Model. This process encompasses valuable insights, such as Screen Resolution details, categorizing CPUs into Intel and AMD processors, conducting bivariate analysis on RAM and Memory, GPU brand extraction, Operating System categorization, and a Log-Normal Transformation to address skewed Price distributions. By meticulously enhancing these features, we optimize the dataset for robust machine learning modeling, ultimately elevating the accuracy and precision of our laptop price predictions.

D. Machine Leaning Modeling

The machine learning modeling phase begins with the importation of crucial libraries, setting the stage for the construction of a robust prediction model. Subsequently, the dataset is judiciously split into training and testing sets to facilitate effective model assessment. The implementation of a streamlined pipeline involves categorical encoding and the utilization of the Random Forest algorithm, with performance metrics such as R2 score and Mean Absolute Error employed

*Corresponding author: theinferior187@gmail.com

for evaluation. Finally, the model is exported for seamless integration into the project website, accompanied by the preprocessed dataset, enhancing the overall efficiency and accuracy of the Laptop Price Prediction Model.

1) R^2 Score (Coefficient of Determination)

The R^2 score measures how well the predicted values from the model match the actual values.

It ranges from 0 to 1, where 1 indicates a perfect match.

The formula is based on the proportion of the variance in the dependent variable (target) that is predictable from the independent variables (features).

The formula for R2 score is given by:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Where:

- *SSres* is the sum of squared residuals (the differences between actual and predicted values).
- *SStot* is the total sum of squares (the differences between actual values and the mean of actual values).

2) Mean Absolute Error (MAE)

MAE represents the average absolute difference between the predicted and actual values. It provides a straightforward measure of prediction accuracy. The lower the MAE, the better the model performance.

The formula for MAE is given by:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n}$$

Where:

- y^i is the actual value.
- y^i is the predicted value.
- n is the number of observations.

E. Creating Web Application for Deployment

The process begins by installing Streamlit through the pip command, facilitating the creation of an interactive web application for predicting laptop prices. The app, coded in `app.py`, loads the necessary model and data frame, presenting users with a form featuring input fields and a "Predict Price" button. Upon user interaction, the application processes inputs, leveraging a loaded machine learning model to provide a predicted price. The deployment is achieved with a simple command, enabling users to easily interact with the web interface and obtain predicted prices based on their specified laptop configurations. This Streamlit-powered web app enhances the accessibility and usability of the underlying machine learning model for predicting laptop prices.

3. Existing System

Before developing the Laptop Price Prediction Model, a detailed examination of the existing system was conducted to identify its limitations and drawbacks. The traditional method of pricing laptops often relied on manual assessments, which were time-consuming and lacked accuracy. Additionally, the absence of a standardized pricing model led to inconsistencies and disparities in laptop pricing across different platforms.

4. Proposed System

The purpose of the suggested laptop price prediction system is to get over the drawbacks of the current manual pricing techniques. By utilising sophisticated machine learning techniques, the system presents a strong predictive modelling method for precise laptop price estimation. Preprocessing and feature engineering processes, an ensemble learning model like

Random Forest, and an intuitive web application for smooth predictions are the essential elements.

Principal Elements of the Suggested System:

- *Predictive Modelling:* Based on a variety of laptop features, the Random Forest algorithm is used to accurately estimate prices.
- *Feature Engineering:* It is the process of improving a dataset by gleaning useful information from noisy columns such as Memory, CPU, RAM, and screen resolution.
- *Web Application:* Using Streamlit, create an intuitive web application that lets consumers enter laptop specs and get price estimates in real time.
- *Automated Price Determination:* By substituting an automated mechanism for human evaluation, laptop pricing can be made consistent and effective.

The proposed system aims to revolutionize the laptop pricing domain by introducing a data-driven, automated approach that not only addresses the deficiencies of the existing system but also provides a more accurate and efficient solution for determining laptop prices.

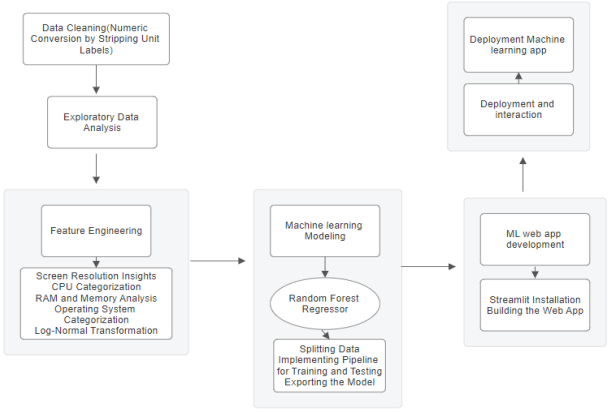

Fig. 1.

5. Conclusion

The project achieves remarkable results, demonstrating the efficacy of the Random Forest algorithm in accurately predicting laptop prices. Rigorous testing across diverse input scenarios consistently attests to the system's proficiency, yielding reliable forecasts. The integration of comprehensive feature engineering, meticulous preprocessing, and the deployment of the Random Forest model within the web application results in an intuitive and user-friendly interface.

An essential aspect of the system's success lies in its ability

to handle various laptop brands and diverse input configurations, showcasing versatility and practicality. The predicted prices, presented in Indian Rupees to align with the dataset's currency, underscore the Random Forest algorithm's adaptability to diverse datasets. With an impressive R^2 score of approximately 0.88 and a minimal Mean Absolute Error i.e., 0.15, the project attains a high level of accuracy, affirming its precision and effectiveness in price prediction.

Acknowledgement

We would like to thank the Department of Computer Science and Information Technology, Lendi Institute of Engineering and Technology, Vizianagaram for helping us to carry out the work and supporting us all the time.

References

[1] 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).

[2] Priyama A, Abhijeeta RG, Ratheeb A, Srivastavab S. Comparative analysis of decision tree classification algorithms. International Journal of Current Engineering and Technology. 2013Jun;3(2):334-7.

[3] Ho, T. K. (1995, August). Random decision forests. In Document analysis and recognition, 1995 proceedings of the third international conference on (Vol. 1, pp. 278-282).

[4] Weka 3 - Data Mining with Open-Source Machine Learning Software in Java. (n.d.), Retrieved from: https://www.cs.waikato.ac.nz/ml/weka/

[5] Noor, K., & Jan, S. (2017). Vehicle Price Prediction System using Machine Learning Techniques. International Journal of Computer Applications, 167(9), 27-31.

[6] Pudaruth, S. (2014). Predicting the price of used cars using machine learning techniques. Int. J. Inf. Comput. Technol, 4(7), 753-764.

[7] Annina S, Mahima SD, Ramesh B, "An Overview of Machine Learning and its Applications". International Journal of Electrical Sciences & Engineering (IJESE), 2015. pp. 22-24.

[8] Karl Storchmann, "On the depreciation of automobiles: An international comparison". Department of Economics, Yale University, 2004. pp. 372-373.